MASTER EN SCIENCES MATHÉMATIQUES

Dualité et stabilité en optimisation non convexe

HEYEZ, Anne-Marie

Award date:
1978

Awarding institution:
Universite de Namur

Année académique :
1977 - 1978

Dualité et Stabilité en optimisation non convexe.

Je remercie Messieurs
Nguyen Van Hien et J. J Strodiot
pour leur aide et leurs conseils.

Heyez
Anne-Marie.

Introduction

Dans ce mémoire, nous étudions le problème d'optimisation suivant :

$$(P) \qquad \inf_{x \in X} f(x) \overset{\text{ndé}}{=} \alpha$$

X y désigne un ensemble quelconque et f est une fonction de X vers $\overline{\mathbb{R}}$.

Nous aborderons ce problème par le biais de la dualité, c'est à dire nous associerons au problème primal (P) un problème dual et nous étudierons les relations existant entre ces deux problèmes.

Le schéma général de construction du dual est le suivant :

- On insère le problème d'optimisation primal dans une famille de problèmes de minimisation perturbés où (P) correspond à une perturbation σ fixée.

- Les valeurs optimales des problèmes perturbés définissent une fonction de perturbation h qui, en fait, mesure la sensibilité du problème (P) à une perturbation donnée. En outre, cette fonction de perturbation prend la valeur optimale α en σ.

– Le problème dual est alors défini comme étant l'enveloppe supérieure en o d'une classe particulière de minorantes de h choisie en fonction du type de problème envisagé.

Nous étudierons également la stabilité du problème, c'est à dire les hypothèses permettant d'éviter une différence entre la valeur optimale du primal et celle du dual.

Dans le chapitre I, nous rappelons les principaux résultats de dualité et de stabilité en optimisation convexe déjà exposés par P. J Laurent ([2], chapitre VII), I. Ekeland et R. Temam ([3], chapitre III) et R.T Rockafellar ([8], part III).
Dans ce cas, (P) est le problème de minimisation d'une fonction convexe f sur un ensemble convexe X.
La fonction de perturbation h est alors, moyennant quelques hypothèses relativement faibles (cfr I.2 page 19), une fonction convexe. Ceci permet de nous limiter, pour la construction du problème dual, aux minorantes affines de h définies par une fonction bilinéaire mettant en dualité l'espace des perturbations et un autre espace Y.
Nous envisageons également un type particulier de problèmes : les programmes mathématiques convexes dont nous étudions la stabilité.

Dans le chapitre II, nous analysons, parallèlement au cas convexe, les résultats de dualité et de stabilité pour le problème général obtenus par E.J Bälder ([1]) et P.O Lindberg ([4]).

Nous suivons essentiellement la thèse de Bälder, la modification principale apportée est le fait que nous avons choisi différemment la classe de minorantes de la fonction de perturbation h.

En effet, comme le problème considéré est tout à fait général la fonction de perturbation peut être quelconque et en employant de nouveau des minorantes affines de h nous risquerions d'obtenir un écart entre la valeur optimale du primal et celle du dual.

Pour résoudre ce problème, Bälder considérait les minorantes de h définies par une fonction c de "type aiguille en σ" (définition voir chapitre II page 46).

Nous avons préféré considérer une fonction c "pointue en σ" (définition voir chapitre II page 43) comme l'avait fait Lindberg ([4]).

(Une fonction de type aiguille en σ n'étant qu'une fonction pointue en σ particulière.)

Nous démontrons que, dans ce cas, les résultats obtenus par Bälder restent toujours valables et en outre, nous obtenons un théorème supplémentaire (théorème II.2), extrait de la thèse de Lindberg ([4]), nous indiquant que le caractère pointu de c en σ est une condition nécessaire pour obtenir des propriétés de dualité similaires aux propriétés du cas convexe.

Dans le chapitre III, nous étudions les programmes mathématiques avec les outils donnés par la théorie de la dualité non convexe.
Il s'avérera que le lagrangien obtenu pour ce type particulier de problèmes est le prototype des lagrangiens généralisés de la méthode de la fonction de pénalisation extérieure dont nous rappelons brièvement l'idée dans l'annexe V.
Nous exposons deux résultats de convergence de cette méthode obtenus par Bälder.

- le premier théorème donne des conditions suffisantes pour qu'une méthode de maximisation séquentielle du lagrangien engendre une suite de points de X dont les valeurs d'adhérence sont des solutions optimales de (P).
(cfr théorème III.2)
- le second théorème concerne les multiplicateurs exacts c'est à dire des multiplicateurs qui permettent, en une seule maximisation du lagrangien, d'atteindre la valeur optimale du problème initial.
Il donne des conditions suffisantes pour que la solution optimale d'une maximisation du lagrangien en un multiplicateur exact soit une solution optimale de (P).
(cfr théorème III.3)

A la suite du théorème III.2, nous exposons un résultat établi précédemment par Rockafellar ([7]) qui est en fait un cas particulier de celui-ci pour un lagrangien donné.

Nous terminons ce mémoire par une application du théorème III.3 à un problème particulier où nous démontrons l'existence d'un multiplicateur exact vérifiant les hypothèses du théorème.

Remarques préliminaires

Dans ce qui suit, $\mathbb{R}$, $\mathbb{R}_+$, $\overline{\mathbb{R}}$ désignent respectivement

- l'ensemble des nombres réels
- l'ensemble des nombres réels positifs
- l'extension de l'ensemble des nombres réels

soit $\mathbb{R} \cup \{-\infty, +\infty\}$

avec la convention :

$$(+\infty) + (-\infty) = (-\infty) + (+\infty) = +\infty .$$

De façon similaire, nous noterons $\mathbb{R}^m$ l'espace euclidien de dimension m, $\mathbb{R}^m_+$ sa partie positive et $\mathbb{R}^m_-$ sa partie négative.

Chaque fois que nous parlerons de $\mathbb{R}^m$, nous le supposerons muni de la topologie induite par la distance euclidienne.

$\left(d(x, y) \overset{\text{déf}}{=} \sqrt{\sum_{i=1}^{m} (x_i - y_i)^2} \quad \forall x = (x_1, \ldots x_m), y = (y_1, \ldots y_m) \in \mathbb{R}^m \right)$

Dans la suite du travail nous appellerons FONCTIONNELLE une fonction définie sur un espace E quelconque qui est à valeurs dans $\overline{\mathbb{R}}$. Nous noterons $\overline{\mathbb{R}}^E$ l'ensemble des fonctionnelles définies sur E.

L'infimum d'une fonctionnelle sur un ensemble vide est posé égal à $(+\infty)$.

Chapitre I : Résultats généraux de la théorie de la dualité en optimisation convexe

I.1. Définitions préliminaires et généralités sur les fonctions convexes

I.1.1. Définition des fonctionnelles convexes

Définition I.1 :

Soient X un espace vectoriel réel et f appartenant à $\overline{\mathbb{R}}^X$

f est CONVEXE sur X si et seulement si :

$$\forall x_1, x_2 \in X \; ; \; \forall \lambda_1, \lambda_2 \in \mathbb{R} \; , \; \lambda_1, \lambda_2 > 0 \text{ et } \lambda_1 + \lambda_2 = 1$$

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

Remarque : Si f est une fonctionnelle définie sur une partie convexe C de X, on peut la prolonger à X tout entier par :

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si } x \in C \\ +\infty & \text{si } x \notin C \end{cases}$$

Nous dirons que f est convexe sur le convexe C si et seulement si $\tilde{f}$ est convexe sur X

<u>Proposition I.1</u>

Soit une fonctionnelle f définie sur le convexe C inclus dans X

f est convexe sur le convexe C si et seulement si :

$$\forall x_1, x_2 \in C \; ; \; \forall \lambda_1, \lambda_2 \in \mathbb{R}, \; \lambda_1, \lambda_2 > 0 \text{ et } \lambda_1 + \lambda_2 = 1$$

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

<u>Démonstration</u>

<u>Condition nécessaire</u> :

Elle est évidente par la définition de f convexe sur le convexe C

<u>Condition suffisante</u>:

Nous devons vérifier que $\tilde{f}$ est convexe sur X.

Soient x_1, x_2 quelconques dans X et λ_1, λ_2 deux nombres réels strictement positifs arbitraires dont la somme vaut un.

Nous devons prouver l'inégalité suivante :

$$\tilde{f}(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 \tilde{f}(x_1) + \lambda_2 \tilde{f}(x_2)$$

Deux cas peuvent se présenter :

1) Les points x_1 et x_2 appartiennent à C

Dans ce cas, $\tilde{f}(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 \tilde{f}(x_1) + \lambda_2 \tilde{f}(x_2)$

car C est convexe et f coïncide avec $\tilde{f}$ sur C

2) L'un des deux points, au moins, n'appartient pas à C

Pour fixer les idées supposons que c'est x_1.

Par conséquent, $\tilde{f}(x_1) = +\infty$

donc, $\tilde{f}(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 \tilde{f}(x_1) + \lambda_2 \tilde{f}(x_2) = +\infty$.

I.1.2. Semi-continuité des fonctionnelles convexes

Soit X un espace vectoriel topologique séparé.

Définitions I2 :

Une fonctionnelle f appartenant à $\overline{\mathbb{R}}^X$ est dite SEMI-CONTINUE INFÉRIEUREMENT (noté s.c.i) en un élément x_0 de X si, pour tout réel k strictement inférieur à $f(x_0)$ il existe un voisinage $\mathcal{U}$ de x_0 tel que :

$$\forall x \in \mathcal{U} \quad k < f(x)$$

Ceci est équivalent à : $f(x_0) = \liminf\limits_{x \to x_0} f(x)$ ([2] page 332)

Une fonctionnelle f appartenant à $\overline{\mathbb{R}}^X$ est dite SEMI-CONTINUE SUPÉRIEUREMENT (noté s.c.s) en un élément x_0 de X si, pour tout réel k strictement supérieur à $f(x_0)$ il existe un voisinage $\mathcal{U}$ de x_0 tel que :

$$\forall x \in \mathcal{U} \quad k > f(x)$$

Ceci est équivalent à : $f(x_0) = \limsup\limits_{x \to x_0} f(x)$

Une fonctionnelle f de $\overline{\mathbb{R}}^X$ est s.c.i (s.c.s) si elle est s.c.i (s.c.s) en tout point x_0 de X.

Remarquons qu'une fonctionnelle f de $\overline{\mathbb{R}}^X$ est continue en x_0 si elle est à la fois s.c.i et s.c.s en x_0.

Proposition I.2

Soit une fonctionnelle f appartenant à $\overline{\mathbb{R}}^X$, alors les trois propositions suivantes sont équivalentes :

1) f est s.c.i
2) Pour tout réel λ, $S_\lambda = \{x \in X \mid f(x) \leq \lambda\}$ est fermé
3) L'épigraphe de f est fermé dans $X \times \mathbb{R}$ muni de la topologie produit.
   où l'épigraphe de $f = \{(x,\lambda) \in X \times \mathbb{R} \mid f(x) \leq \lambda\}$

(démonstration voir [2] pages 330-331)

Des résultats similaires existent pour f s.c.s ([2] page 333)

I.1.3. Polaire d'une fonctionnelle

Définition I.3

On dit que deux espaces vectoriels réels X et Y sont mis en DUALITÉ par une fonctionnelle bilinéaire notée

$$<.,.> : X \times Y \longrightarrow \mathbb{R} : (x,y) \longmapsto <x,y>$$

si :

$$\begin{cases} \forall x \in X,\ x \neq 0_X,\ \exists y \in Y \text{ tel que } <x,y> \neq 0 \\ \forall y \in Y,\ x \neq 0_Y,\ \exists x \in X \text{ tel que } <x,y> \neq 0 \end{cases}$$

où 0_X et 0_Y désignent le neutre pour l'addition, respectivement de X et de Y.

Exemple 1 : Posons $X = Y = \mathbb{R}^m$ et $<x,y> = \sum_{i=1}^{m} x_i \cdot y_i$ le produit scalaire usuel dans $\mathbb{R}^m$

Dans ce cas, la définition est vérifiée

Définition I.4

Une topologie localement convexe sur un espace vectoriel réel topologique X est dite COMPATIBLE avec la DUALITÉ entre X et Y si la fonctionnelle linéaire, qui, à x appartenant à X, associe le réel $\langle x, y \rangle$, est continue pour tout élément y de Y et si toute fonctionnelle linéaire continue sur X est représentable de cette façon.

Exemple 2 : Reprenons les hypothèses de l'exemple 1.
Dans ce cas, la topologie engendrée par la métrique euclidienne est compatible avec la dualité définie par le produit scalaire.
(Ceci est une application du théorème de représentation dans les espaces de Hilbert)

Enoncé du théorème de représentation :

Soit H un espace de Hilbert dont le produit scalaire est noté $\langle .,. \rangle$ et H' l'espace des formes linéaires continues sur H.
Nous avons que :
i) $\forall f \in H'$, $\exists y \in H$ tel que $f(x) = \langle x, y \rangle$ $\forall x \in H$
ii) $\forall y \in H'$, $\exists f \in H$ tel que $f(x) = \langle x, y \rangle$ $\forall x \in H$

Dans la suite du chapitre nous désignerons par X et Y deux espaces vectoriels topologiques localement convexes (E.V.T.L.C) et par $\langle .,. \rangle$ la fonctionnelle bilinéaire mettant X et Y en dualité.
Les topologies sur X et sur Y sont supposées compatibles avec la dualité.

Définition I.5

Etant donné une fonctionnelle f appartenant à $\overline{\mathbb{R}}^X$, on définit la POLAIRE de f notée f^* par :

$$f^*(y) = \sup_{x \in X} (\langle x, y\rangle - f(x)) \quad \forall y \in Y$$

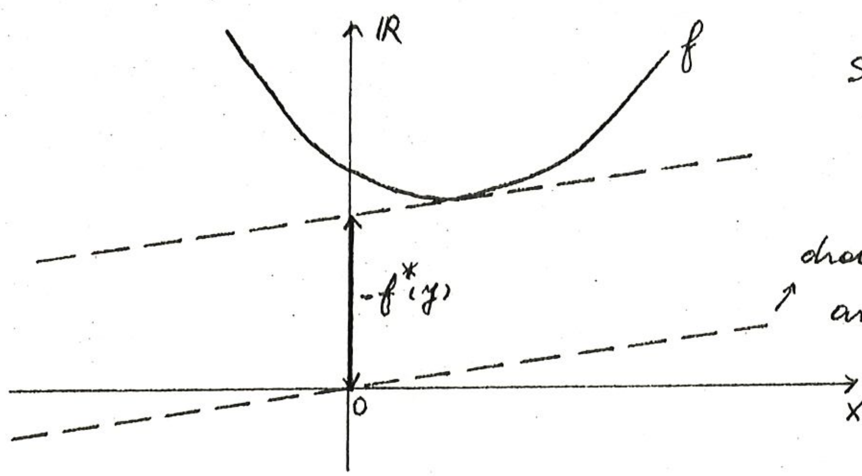
En transformant l'expression de f^* nous pouvons donner une interprétation géométrique intéressante à la polaire d'une fonctionnelle.

$$f^*(y) = - \inf_{x \in X} (f(x) - \langle x, y\rangle) \quad \forall y \in Y$$

en exprimant que l'infimum est le plus grand des minorants, nous obtenons :

$$f^*(y) = - \sup \{\eta \mid \eta \in \mathbb{R} \,;\, \langle x, y\rangle + \eta \leq f(x) \ \forall x \in X\} \quad \forall y \in Y$$

En fait, la fonctionnelle f^* résume les minorantes affines de f. A une direction y, elle associe $f^*(y) = -\eta \in \overline{\mathbb{R}}$ où η est la valeur maximale prise à l'origine de X par une minorante affine de f de direction y.
(par minorante affine de direction y nous désignons toute fonctionnelle du type $\langle \cdot, y\rangle + \eta$ qui minore f $(\eta \in \mathbb{R})$)



Soit $X = Y = \mathbb{R}$ et $\langle x, y\rangle = x \cdot y$

droite de coefficient angulaire y.

On définit de même f^{**} appartenant à $\overline{\mathbb{R}}^X$ comme étant la polaire de f^*.

$$f^{**}(x) \overset{\text{déf}}{=} \sup_{y \in Y} (\langle x, y \rangle - f^*(y)).$$

Géométriquement, f^{**} correspond exactement à l'enveloppe supérieure des minorantes affines de f.

En effet, soient un élément y de Y et un réel η tels que :

$$\langle x, y \rangle + \eta \leq f(x), \forall x \in X$$

Ceci est équivalent à :

$$\sup_{x \in X} \{\langle x, y \rangle - f(x)\} \leq - \eta$$

donc à :

$$f^*(y) \leq - \eta$$

Or, $f^{**}(y) \overset{\text{déf}}{=} \sup_{y \in Y} (\langle x, y \rangle - f^*(y))$

Par conséquent,

$$f^{**}(y) = \sup_{\substack{y \in Y \\ \eta \in \mathbb{R} \mid \eta \leq - f^*(y)}} (\langle x, y \rangle + \eta)$$

Ce qui revient, par la première étape à

$$f^{**}(y) = \sup_{\substack{y \in Y \\ \eta \in \mathbb{R} \\ \langle x, y \rangle + \eta \leq f(x), \forall x \in X}} (\langle x, y \rangle + \eta)$$

Définition I.6

Nous noterons $\Gamma(X)$ l'ensemble des fonctionnelles de $\overline{\mathbb{R}}^X$ qui s'expriment comme l'enveloppe supérieure d'une famille de fonctionnelles affines continues c'est à dire :

$$\Gamma(X) = \{ f \in \overline{\mathbb{R}}^X \mid f(x) = \sup_{i \in I} (\langle x, y_i \rangle - r_i) \}$$

où I est une famille quelconque d'indices et pour tout i dans I, y_i appartient à Y et r_i est un réel.

Il est évident que les fonctionnelles ω_1 et ω_2, définies par : $\omega_1(x) = +\infty$ et $\omega_2(x) = -\infty, \forall x \in X$, sont des éléments de $\Gamma(X)$.
(il suffit, pour ω_1, de considérer la famille de toutes les fonctionnelles affines continues et pour ω_2, on prend I égal à l'ensemble vide.)

Nous noterons $\Gamma_0(X)$ l'ensemble $\Gamma(X) \setminus \{\omega_1, \omega_2\}$.

Remarque:

$$\forall f \in \overline{\mathbb{R}}^X , \quad f^* \in \Gamma(Y) \text{ et } f^{**} \in \Gamma(X).$$

Définition I.7

Une fonctionnelle f appartenant à $\overline{\mathbb{R}}^X$ est dite PROPRE si elle ne prend pas la valeur $(+\infty)$ identiquement et si elle ne prend jamais la valeur $(-\infty)$.

Théorème I.1

L'ensemble $\Gamma_0(X)$ coïncide avec l'ensemble des fonctionnelles de $\overline{\mathbb{R}}^X$ qui sont convexes, s.c.i et propres.

(démonstration, voir [2] page 340)

Théorème I.2

Soit f appartenant à $\overline{\mathbb{R}}^X$.
f et f^{**} coïncident si et seulement si f appartient à $\Gamma(X)$.

(démonstration, voir [2] page 344)

I.1.4. Sous différentiel d'une fonctionnelle

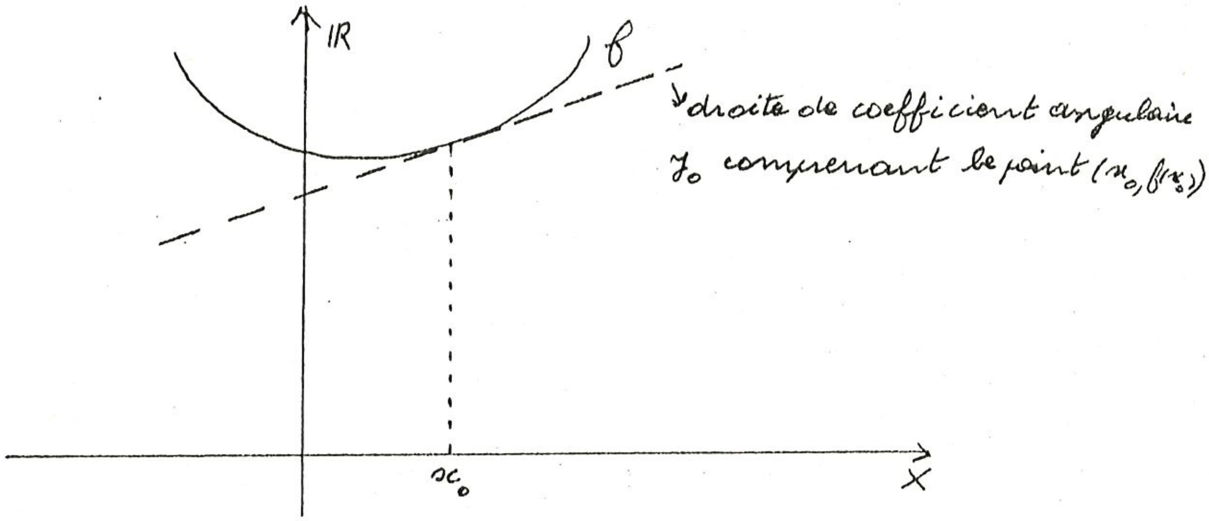
Définitions I.8

Soit f appartenant à $\overline{\mathbb{R}}^X$, on dit qu'un élément y_0 de Y est un SOUS-GRADIENT de f au point x_0 de X si :

$$\left| \begin{array}{l} f(x_0) \text{ est fini} \\ f(x) \geq f(x_0) + \langle x - x_0, y_0 \rangle \quad \forall x \in X \end{array} \right.$$

En d'autres termes, y_0 est un sous-gradient de f au point x_0 si il existe une minorante affine continue de f, de pente y_0 et qui prend en x_0 la valeur finie $f(x_0)$.

Posons $X = Y = \mathbb{R}$ et $\langle x, y \rangle = x \cdot y \quad x \in X,\ y \in Y$.



Dans le cas particulier envisagé sur le dessin, y_0 est un sous gradient de f au point x_0.

L'ensemble (éventuellement vide) des sous gradients de f en x_0 sera appelé SOUS-DIFFÉRENTIEL de f en x_0 et noté $\partial f(x_0)$.

Dans la suite de ce travail, nous désignerons par minorante affine de f EXACTE en x_0, une minorante affine continue qui prend en x_0 la valeur $f(x_0)$

Propriétés I.1

1) L'élément y_0 de Y est un sous-gradient de la fonctionnelle f appartenant à $\overline{\mathbb{R}}^X$ en $x_0 \in X$ si et seulement si :

$$f(x_0) + f^*(y_0) = \langle x_0, y_0 \rangle$$

(démonstration, [2] page 350).

2) Soient f appartenant à $\overline{\mathbb{R}}^X$ et x_0 un élément de X

(i) Si $\partial f(x_0) \neq \emptyset$ alors, $f^{**}(x_0) = f(x_0)$

(ii) Si $f^{**}(x_0) = f(x_0)$ alors $\partial f^{**}(x_0) = \partial f(x_0)$

(iii) Si $y_0 \in \partial f(x_0)$ alors $x_0 \in \partial f^*(y_0)$

(démonstration, [2] page 351)

3) Soit f appartenant à $\overline{\mathbb{R}}^X$.

Un élément $\bar{x}$ de X vérifie

$$f(\bar{x}) = \min_{x \in X} f(x) \qquad (\text{avec } f(\bar{x}) \text{ fini}) \qquad (1)$$

si et seulement si 0_Y appartient à $\partial f(\bar{x})$

En outre, si f appartient à $\Gamma_0(X)$, l'ensemble des éléments $\bar{x}$ de X qui vérifient l'égalité (1) est égal à $\partial f^*(0_Y)$

(démonstration, [2] page 352)

Commentaire :

Les propriétés 2)(i), (ii) et 3) se justifient aisément d'un point de vue géométrique.

En effet, 2)(i) signifie simplement que, si f admet une minorante affine exacte en x_0 alors, en ce point, elle coïncide avec l'enveloppe supérieure de ses minorantes affines.

2)(ii) traduit le fait que, si f et l'enveloppe supérieure de ses minorantes affines (f^{**}) coïncident en x_0 alors toute minorante affine de f exacte en x_0 est une minorante affine de f^{**} exacte en x_0.

3) nous dit que f admet une minorante affine constante (égale à $f(\bar{x})$ pour tout x dans X) ssi $\bar{x}$ est un minimum de f sur X.

I.2. Stabilité et dualité en optimisation convexe

I.2.1. Perturbation d'un problème de minimisation et problème dual associé

Soient X et Y deux E.V.T.L.C en dualité par rapport à la fonctionnelle bilinéaire $<x, y>$, $x \in X$, $y \in Y$.

Considérons le problème de minimisation suivant.

$$(P) \qquad \inf_{x \in X} f(x) \overset{\text{ndé}}{=} \alpha$$

où f est une fonctionnelle convexe sur X.

Remarque:

Le problème (P) comprend les problèmes de minimisation du type:

$$\inf_{x \in C} f(x)$$

où C est un convexe inclus dans X et f est une fonctionnelle convexe sur le convexe C.
En effet, nous pouvons considérer le prolongement $\tilde{f}$ de f sur X qui est une fonctionnelle convexe sur X (cfr proposition I.1) et qui, en outre, vérifie:

$$\inf_{x \in C} f(x) = \inf_{x \in X} \tilde{f}(x)$$

Considérons maintenant U et V deux E.V.T.L.C en dualité par rapport à la fonctionnelle bilinéaire notée $[u, v]$, $u \in U$, $v \in V$.

L'espace U sera appelé l'espace des PERTURBATIONS pour le problème (P).

Supposons donnée une fonctionnelle φ convexe, s.c.i, propre définie sur $X \times U$ telle que :

$$f(x) = \varphi(x, 0) \qquad \forall x \in X$$

où 0 désigne le neutre pour l'addition dans U.

Nous associons à une perturbation u de U le problème perturbé suivant :

$$(P_u) \qquad \inf_{x \in X} \varphi(x, u) \overset{\text{noté}}{=} h(u)$$

Le problème initial (P) correspond à la valeur $u = 0$ et par conséquent nous avons :

$$\alpha = h(0)$$

h est donc une fonctionnelle qui mesure la variation de la valeur optimale de (P) pour une perturbation u donnée.

Le choix de φ implique de façon immédiate que h est convexe sur U.

En effet, soient u_1, u_2 quelconques dans U et λ un réel arbitraire tel que : $0 < \lambda < 1$, nous avons :

$$h(\lambda u_1 + (1-\lambda) u_2) \overset{\text{déf}}{=} \inf_{x \in X} \varphi(x, \lambda u_1 + (1-\lambda) u_2)$$

comme X est un espace vectoriel,

$\forall x \in X \quad \exists x_1, x_2 \in X$ tels que $\quad x = \lambda x_1 + (1-\lambda) x_2$

donc,

$$h(\lambda u_1 + (1-\lambda) u_2) = \inf_{x_1, x_2 \in X} \varphi(\lambda x_1 + (1-\lambda) x_2 \,,\, \lambda u_1 + (1-\lambda) u_2)$$

$$\leq \inf_{x_1, x_2 \in X} [\lambda \varphi(x_1, u_1) + (1-\lambda) \varphi(x_2, u_2)]$$

(car φ est convexe sur $X \times U$)

Il s'ensuit :

$$h(\lambda u_1 + (1-\lambda) u_2) \leq \lambda h(u_1) + (1-\lambda) h(u_2)$$

La fonction de perturbation h est donc convexe sur U mais, en général, elle n'appartient pas à $\Gamma_0(U)$.

Définissons maintenant le problème dual associé à (P)

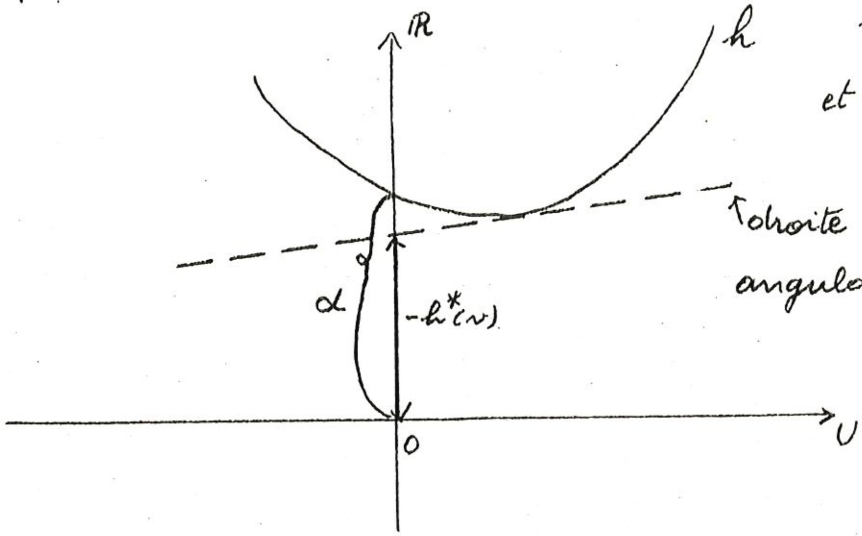
Posons $g(v) = h^*(v) \quad \forall v \in V$; par définition de h^*, g appartient à $\Gamma(V)$

Considérons le problème suivant :

$$(Q) \qquad \inf_{v \in V} g(v) \overset{\text{noté}}{=} \beta$$

Le problème (Q) sera appelé le problème dual de (P) par rapport à la famille des problèmes perturbés (P_u) qui est entièrement définie par la fonctionnelle φ.

Le problème dual consiste, au signe près, à chercher la valeur maximale prise par une minorante affine de h en σ.



Posons $U = V = \mathbb{R}$ et $[u, v] = u \cdot v$

Remarquons que, comme h^{**} est l'enveloppe supérieure des minorantes affines de h nous avons :

$$-\alpha = -h(0) \leq -h^{**}(0) = \beta$$

Remarque :

L'hypothèse "h appartient à $\Gamma(X)$" est une condition suffisante d'égalité entre $-\alpha$ et β. (cfr théorème I.2)

Appelons A l'ensemble des solutions de (P) et B l'ensemble des solutions de (Q).

Proposition I.3

Si f appartient à $\Gamma(X)$ et si α, β sont finis alors $A = \{\bar{x} \in X \mid \alpha = f(\bar{x})\} = \partial f^*(0)$

et $B = \{\bar{v} \in V \mid \beta = g(\bar{v})\} = \partial g^*(0)$

Cette proposition est une application immédiate de la propriété 3 sur les sous-différentiels.

Remarque :

Nous pouvons à partir du dual recommencer une construction similaire à celle réalisée sur le primal.

Remarquons d'abord que les espaces $X \times U$ et $Y \times V$ sont deux E.V.T.L.C (nous les supposons munis de la topologie produit) en dualité par rapport à la fonctionnelle bilinéaire :

$$<(x,u);(y,v)> \overset{déf}{=} <x,y> + [u,v]$$

Soit Ψ la polaire de la fonctionnelle φ. Ψ appartient donc à $\Gamma(Y \times V)$ et comme φ est propre, Ψ est par conséquent convexe, s.c.i, propre sur $Y \times V$

$$\Psi((y,v)) \overset{déf}{=} \sup_{\substack{x \in X \\ u \in U}} (<x,y> + [u,v] - \varphi(x,u))$$

$$\text{Or, } h^*(v) \overset{déf}{=} \sup_{u \in U} ([u,v] - \inf_{x \in X} \varphi(x,u)) \qquad \forall v \in V$$

$$= \sup_{\substack{x \in X \\ u \in U}} (<x,0> + [u,v] - \varphi(x,u)) \qquad \forall v \in V$$

finalement, nous obtenons :

$$h^*(v) = \Psi(0,v) \qquad \forall v \in V$$

En considérant maintenant Y comme l'espace des des perturbations, nous définissons pour une perturbation y de Y le problème dual perturbé.

$$(Q_y) \qquad \inf_{v \in V} \Psi(y,v) \overset{\text{noté}}{=} k(y)$$

Le problème dual (Q) correspond à la perturbation $y = 0_Y$ et donc,

$$\beta = k(0_Y)$$

Nous remarquons de nouveau que la fonctionnelle k est convexe sur Y mais, en général, elle n'appartient pas à $\Gamma_0(Y)$

Notons que la construction réalisée ci-dessus est complètement symétrique par rapport à (P) et à (Q). Nous avons :

$$-\beta = -k(0_Y) \leq -k^{**}(0_Y) = \operatorname*{Inf}_{x \in X} k^*(x)$$

or, un calcul similaire à celui réalisé précédemment pour h^* nous montre que :

$$k^*(x) = \Phi(x,0) = f(x) \qquad \forall x \in X$$

et par conséquent,

$$-\beta = -k(0) \leq -k^{**}(0) = \alpha$$

Donc, le problème (P) est le dual de (Q) par rapport à la famille des problèmes perturbés (Q_y) laquelle est entièrement définie par la fonctionnelle Φ.

I.2.2. Stabilité d'un problème de minimisation convexe

Définitions I.9

Le problème (P) sera dit STABLE si $h(0)$ est fini et si h est continue en $0 \in U$.

Le problème (P) sera dit INF-STABLE si $h(0)$ est fini et si h est s.c.i en $0 \in U$ c'est à dire si :

$$h(0) = \liminf_{u \to 0} h(u)$$

Théorème I.3

Les trois propositions suivantes sont équivalentes :

(i) le problème (P) est inf-stable

(ii) le problème (Q) est inf-stable

(iii) $-\beta = \alpha$ = un nombre fini

(démonstration voir [2], page 396)

La semi-continuité inférieure de la fonction de perturbation en 0 est donc une condition nécessaire et suffisante pour obtenir, au signe près, l'égalité entre la valeur optimale du problème primal (P) et la valeur optimale du problème dual (Q).

Définition I.10

Nous appellerons LAGRANGIEN de (P) la fonctionnelle suivante définie sur le produit $X \times V$.

$$\ell(x, v) \overset{\text{déf}}{=} \sup_{u \in U} \left([u, v] - \varphi(x, u) \right)$$

Pour un x fixé, la fonctionnelle $l_x : V \to \bar{\mathbb{R}}$ définie par $l_x(v) = l(x,v)$ coïncide avec φ_x^* la polaire de la fonctionnelle φ_x définie par $\varphi_x(u) = \varphi(x,u)$ $\forall u \in U$.

Donc, pour un x fixé, l_x appartient à $\Gamma(V)$

En outre, $\varphi_x \in \Gamma(U)$ et par conséquent :

$$\varphi(x,u) = \varphi_x(u) = \varphi_x^{**}(u) = \sup_{v \in V} ([u,v] - l(x,v))$$

Il s'ensuit, pour $u = 0$, que :

$$\alpha = \inf_{x \in X} \sup_{v \in V} (-l(x,v))$$

ou

$$-\alpha = \operatorname{Sup}_{x \in X} \operatorname{Inf}_{v \in V} l(x,v)$$

En comparant la définition de ψ et de l nous obtenons également :

$$\psi(y,v) = \operatorname{Sup}_{x \in X} (\langle x,y \rangle + l(x,v)) \quad \forall y \in Y, \forall v \in V$$

donc, pour $y = 0_Y$, cela implique :

$$\beta = \inf_{v \in V} \operatorname{Sup}_{x \in X} l(x,v)$$

Ainsi, l'égalité $-\beta = \alpha$ correspond à la possibilité de pouvoir échanger $\operatorname{Inf}_{v \in V}$ et $\sup_{x \in X}$ devant le lagrangien

Définition I.11

Nous dirons qu'un couple $(\bar{x}, \bar{v})$ appartenant à $X \times V$ est un POINT de SELLE d'une fonctionnelle L de $\bar{\mathbb{R}}^{X \times V}$ si :

$$L(\bar{x}, v) \leq L(\bar{x}, \bar{v}) \leq L(x, \bar{v}) \qquad \forall x \in X, \forall v \in V$$

Remarquons que si $(\bar{x}, \bar{v})$ appartenant à $X \times V$ est un point de selle d'une fonctionnelle L de $\bar{\mathbb{R}}^{X \times V}$ alors $\bar{x}$ minimise $L(., \bar{v})$ et $\bar{v}$ maximise $L(\bar{x}, .)$

Théorème I.4

Si $(\bar{x}, \bar{v})$ est un point de selle du lagrangien $(\bar{x} \in X, \bar{v} \in V)$ alors $\bar{x}$ est une solution optimale de (P) et $\bar{v}$ est une solution optimale de (Q)

(démonstration voir [2] page 402)

I.3. Problème particulier de minimisation convexe

I.3.1. Définition du problème

Définition I. 12

Soit le problème de minimisation suivant :

$$(P) \qquad \inf_{x \in X} f(x) = \alpha$$

X étant supposé muni d'une topologie.

Nous dirons que (P) est un PROGRAMME MATHÉMATIQUE. si il existe une fonctionnelle f_0 appartenant à $\overline{\mathbb{R}}^X$ et une application $G : X \to Z$ où Z est un espace vectoriel topologique ordonné par un cône convexe fermé C dans Z

$(z' \leq z \text{ ssi } z' - z \in C \quad \forall z, z' \in Z)$

telles que :

$$\begin{cases} f(x) = f_0(x) & \text{pour } x \in X \text{ tel que } G(x) \leq 0_Z \quad (G(x) \in C) \\ f(x) = +\infty & \text{pour } x \in X \text{ tels que } G(x) \not\leq 0_Z \quad (G(x) \notin C) \end{cases}$$

Remarques :

1) 0_Z désigne le neutre pour l'addition dans Z.
2) Si nous prenons $Z = \mathbb{R}$, nous obtenons l'ordre habituel sur $\mathbb{R}$ en posant $C = \mathbb{R}_-$.

Dans ce qui suit, nous prendrons Z égal à U et donc 0_Z égal à 0.

Plaçons-nous dans le cas convexe c'est à dire dans le cas particulier où :

- X est un E.V.T.L.C
- f_0 est une fonctionnelle convexe sur X
- G est convexe par rapport à l'ordre défini dans U

ce qui signifie :

$\forall x, x' \in X \quad , \quad \forall \lambda \in]0,1[$

$G(\lambda x + (1-\lambda)x') - \lambda G(x) - (1-\lambda) G(x') \in C$

Considérons la famille des problèmes perturbés $\{(P_u) \mid u \in U\}$ avec :

$$(P_u) \qquad \inf_{x \in X} \{ f_0(x) \mid G(x) \leq u \} = h(u) \qquad \forall u \in U$$

Ceci revient à définir φ appartenant à $\bar{\mathbb{R}}^{X \times U}$ par :

$$\varphi(x,u) = \begin{cases} f_0(x) & \text{si } \; G(x) \leq u \\ +\infty & \text{sinon} \end{cases} \qquad \forall x \in X, \forall u \in U$$

Vérifions que la fonction φ est bien choisie :

1) $\varphi(x,0) = f(x) \qquad \forall x \in X$

2) φ est une fonction convexe sur $X \times U$

Preuve :

Soient (x_1,u_1), (x_2,u_2) quelconques dans $X \times U$ et λ un réel arbitraire appartenant à l'intervalle $]0,1[$.

Nous devons vérifier :

$$\varphi(\lambda x_1+(1-\lambda)x_2\,,\,\lambda u_1+(1-\lambda)u_2) \leq \lambda\varphi(x_1,u_1)+(1-\lambda)\varphi(x_2,u_2) \qquad (1)$$

Nous envisageons deux cas:

<u>1^{er} cas</u>: $G(\lambda x_1+(1-\lambda)x_2) \leq \lambda u_1+(1-\lambda)u_2$

Ceci implique :

$$\varphi(\lambda x_1+(1-\lambda)x_2\,,\,\lambda u_1+(1-\lambda)u_2) = f_0(\lambda x_1+(1-\lambda)x_2)$$

or, f_0 est convexe et par conséquent,

$$\varphi(\lambda x_1+(1-\lambda)x_2\,,\,\lambda u_1+(1-\lambda)u_2) \leq \lambda f_0(x_1)+(1-\lambda)f_0(x_2) \qquad (2)$$

a) Si $G(x_1)\leq u_1$ et $G(x_2)\leq u_2$, nous avons immédiatement l'inégalité (1) (cfr (2) et définition de φ)

b) Si $G(x_1)\nleq u_1$ ou $G(x_2)\nleq u_2$ ou les deux; Supposons pour fixer les idées que $G(x_1)\nleq u_1$, il s'ensuit que $\varphi(x_1,u_1)=+\infty$.
Comme $(+\infty)$ est un absorbant suite à la convention initiale (donnée dans les remarques préliminaires) nous aurons :

$$\lambda\varphi(x_1,u_1)+(1-\lambda)\varphi(x_2,u_2) = +\infty$$

et par conséquent l'inégalité (1) est de nouveau vérifiée.

<u>2^{ème} cas</u>: $G(\lambda x_1+(1-\lambda)x_2) \nleq \lambda u_1+(1-\lambda)u_2$

Il s'ensuit :

$$\varphi(\lambda x_1+(1-\lambda)x_2\,,\,\lambda u_1+(1-\lambda)u_2) = +\infty$$

a) Si $G(x_1)\nleq u_1$ ou $G(x_2)\nleq u_2$ ou les deux nous aurons d'office l'inégalité (1).

b) Supposons donc que $G(x_1) \leq u_1$ et $G(x_2) \leq u_2$.
Nous allons démontrer que cette hypothèse est irréalisable dans le cas considéré.

$G(x_1) \leq u_1$ ssi $G(x_1) - u_1 \in C$
$G(x_2) \leq u_2$ ssi $G(x_2) - u_2 \in C$

Donc, puisque C est convexe,
$\lambda G(x_1) - \lambda u_1 + (1-\lambda) G(x_2) - (1-\lambda) u_2 \in C$
Or, par l'hypothèse de convexité sur G
$G(\lambda x_1 + (1-\lambda) x_2) - (\lambda G(x_1) + (1-\lambda) G(x_2)) \in C$

Par conséquent,
$G(\lambda x_1 + (1-\lambda) x_2) - (\lambda u_1 + (1-\lambda) u_2) \in C$ (3)
(cfr C est un cône convexe donc la somme de deux éléments de C appartient à C)

(3) est en contradiction avec l'hypothèse du b) qui est donc irréalisable.

Il s'ensuit que φ est convexe sur $X \times U$ puisque dans tous les cas possibles l'inégalité (1) est vérifiée.

I.3.2. Stabilité du problème

Théorème I.5

Dans le cas d'un programme mathématique convexe, une condition suffisante de stabilité du problème est l'existence d'un élément $\bar{x}$ de X tel que $G(\bar{x})$ appartient à l'intérieur de C et $f_0(\bar{x})$ est différent de $(+\infty)$.
(Cette condition est appelée condition de SLATER)

Démonstration

Nous devons démontrer que h est finie et continue en 0.

Puisque φ est convexe, nous savons déjà que h est convexe. Donc, si nous prouvons que h est bornée supérieurement dans un voisinage de 0, h sera continue en 0 et finie en 0 (cfr [3], page 11, lemme 2.1)

$G(\bar{x})$ appartient à l'intérieur de C, il existe donc un voisinage V de $G(\bar{x})$ tel que V est inclus dans C.

Considérons $W = -V + G(\bar{x})$
Comme nous travaillons dans un espace vectoriel topologique W est un voisinage de 0.
En outre, pour tout u élément de W, $\varphi(\bar{x}, u) = f_0(\bar{x})$ (1)
En effet, $\forall u \in W$, $\exists u_1 \in V$ tel que $u = -u_1 + G(\bar{x})$
donc $G(\bar{x}) - u = u_1 \in V \subset C$
par conséquent $G(\bar{x}) \leq u \quad \forall u \in W$
On en déduit l'égalité (1).

Par définition, $h(u) = \inf_{x \in X} \varphi(x, u) \quad \forall u \in U$

Il s'ensuit :

$\forall u \in W \quad h(u) \leq \varphi(\bar{x}, u) = f_0(\bar{x})$

D'où, h est bornée supérieurement dans un voisinage de 0.

Chapitre II : Généralisation des résultats de dualité et de stabilité aux problèmes d'optimisation non convexes

II.1. Motivations

Dans ce chapitre, nous allons présenter une théorie similaire à celle de la dualité en optimisation convexe mais qui s'applique à des fonctionnelles quelconques.

Soient X un ensemble non vide et f une fonctionnelle définie sur X non identiquement égale à $(+\infty)$. Considérons le problème de minimisation suivant :

$$(P) \qquad \inf_{x \in X} f(x) \overset{\text{noté}}{=} \alpha$$

Remarque :

Le problème (P) comprend les problèmes contraints puisque l'ensemble X et la fonctionnelle f sont quelconques. Cependant, il est utile de remarquer que nous pouvons toujours étendre f à un univers $\tilde{X}$ comprenant X en posant f égale à $(+\infty)$ en dehors de X.

Comme dans le cas convexe (cfr I.2.1) nous allons perturber le problème initial (P).

Soient U un ensemble et σ un élément fixé arbitrairement dans U.

Supposons l'existence d'une fonctionnelle ϕ définie sur $X \times U$ et vérifiant :

$$\phi(x, \sigma) = f(x) \qquad \forall x \in X$$

Le problème (P) peut alors être considéré comme un élément d'une famille de problèmes de minimisation perturbés :

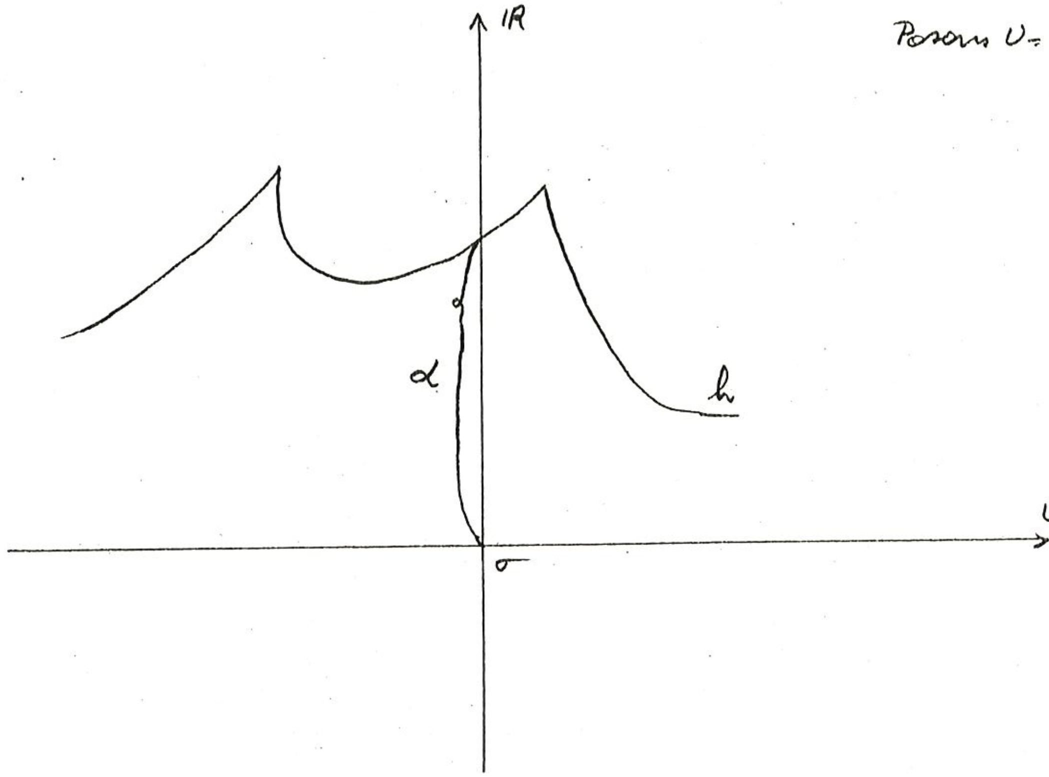
$$(P_u) \qquad \inf_{x \in X} \phi(x, u) = h(u) \qquad u \in U$$

(P) correspond à $u = \sigma$ et donc,

$$\alpha = h(\sigma)$$

Comme pour le cas convexe, h mesure l'influence d'une perturbation u sur la valeur optimale de (P) mais, dans le cas présent, h est en général une fonctionnelle non convexe.
Ceci motive l'emploi d'une autre classe de minorantes pour approcher la valeur de h en σ ; les minorantes affines n'étant, en général, plus assez puissantes comme nous le voyons sur le dessin ci-dessous.

Posons $U = \mathbb{R}$



En fait, nous pouvons déjà presentir que nous serons obligés de considérer des fonctionnelles minorantes pouvant devenir aussi pointues que l'on veut en σ

II.2. Conjuguée d'une fonctionnelle

Dans ce paragraphe, nous allons présenter toutes les notions de base, définitions et propriétés, nécessaires à l'établissement des relations de dualité en optimisation non convexe.
Ce ne sont en réalité que des généralisations des notions introduites dans le chapitre I.

Soient V, un ensemble, et c une fonctionnelle définie sur $U \times V$ et à valeurs dans $\mathbb{R}$.
(c va jouer un rôle analogue à celui de la fonctionnelle bilinéaire qui mettait en dualité les deux espaces vectoriels topologiques dans le problème de minimisation convexe.)

Terminologie

. Toute fonctionnelle appartenant à $\mathbb{R}^{U \times V}$ est appelée fonctionnelle COUPLANTE de U et V.

Nous désignerons par c-FONCTIONNELLE ÉLÉMENTAIRE sur U toute fonctionnelle définie sur U de la forme :

$$c(\cdot, v) + \eta \qquad v \in V, \ \eta \in \overline{\mathbb{R}}$$

Dans le cas où η appartient à $\mathbb{R}$, nous parlerons de c-fonctionnelle élémentaire FINIE.

Nous appellerons C-MINORANTE AFFINE d'une fonctionnelle h appartenant à $\overline{\mathbb{R}}^U$, une c-fonctionnelle élémentaire sur U qui minore h.

L'ensemble des fonctionnelles de $\overline{\mathbb{R}}^U$ s'exprimant comme l'enveloppe supérieure d'une famille de c-fonctionnelles élémentaires sur U est noté $\Gamma^c(U)$

$$a \in \Gamma^c(U) \overset{\text{déf}}{\iff} a(u) = \sup_{i \in I} \left(c(u, v_i) + \eta_i\right), \ \forall u \in U$$

où I est un ensemble arbitraire d'indices et,
$\forall i \in I \quad v_i \in V, \eta_i \in \overline{\mathbb{R}}$.

$\Gamma^c(U)$ est donc l'analogue de $\Gamma(U)$ (I.1.3 définition I.6)

Nous pouvons définir de façon similaire les c-fonctionnelles élémentaires sur V et l'ensemble $\Gamma^c(V)$.

<u>Définitions II.1</u>

Etant donné une fonctionnelle a appartenant à $\overline{\mathbb{R}}^U$, nous définissons la fonctionnelle C-CONJUGUÉE de a, notée a^c et appartenant à $\Gamma^c(V)$, par :

$$a^c(v) \overset{\text{déf}}{=} \sup_{u \in U} \left(c(u, v) - a(u)\right) \qquad \forall v \in V$$

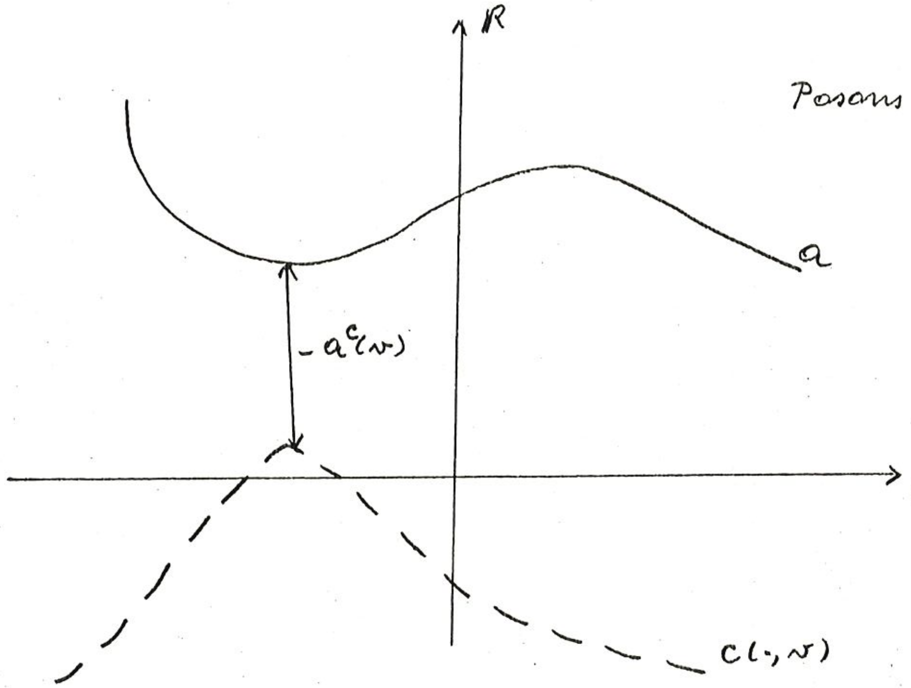
En transformant l'expression de a^c nous pouvons donner à cette fonctionnelle une interprétation géométrique

assez proche de celle de la polaire d'une fonctionnelle.
(cfr I.1.3, définition I.5)

$$a^c(v) = -\operatorname{Sup}\{\eta \mid \eta \in \mathbb{R} ; \eta + c(u,v) \leq a(u) , \forall u \in U\}$$

a^c résume les c-minorantes affines de a. A un élément v de V elle associe $a^c(v) = -\eta \in \overline{\mathbb{R}}$ où η est la valeur de la plus grande translation de $c(\cdot, v)$ telle que $c(\cdot, v) + \eta$ minore la fonctionnelle a.

Posons $U = \mathbb{R}$



Nous définissons de même la fonctionnelle c-conjuguée seconde, notée a^{cc} et appartenant à $\Gamma^c(U)$, par :

$$a^{cc} = (a^c)^c$$

Géométriquement, a^{cc} correspond à l'enveloppe supérieure des c-minorantes affines de a.

En effet,
$$a^{cc}(u) = \sup_{\substack{v \in V \\ \eta \in \mathbb{R} \\ c(u,v)+\eta \leq a(u),\ \forall u \in U}} \left(c(u,v) + \eta \right) \qquad (1)$$

La démonstration de la formule (1) est tout à fait semblable à celle réalisée dans le cadre du cas convexe pour f^{**} (cfr I.1.3, définition I.5)

Propriétés des fonctionnelles c-conjuguées

Soit a une fonctionnelle définie sur U.

1) $a(u) + a^c(v) \geq c(u,v) \qquad \forall u \in U, \forall v \in V$
   (inégalité de Young)
2) $a(u) \geq a^{cc}(u) \qquad \forall u \in U$
3) $a = a^{cc}$ si et seulement si $a \in \Gamma^c(U)$
4) Soit $b \in \Gamma^c(U)$
   Si $a^{cc} \leq b \leq a$ alors $a^{cc} = b$

Commentaires:

La propriété 2) exprime simplement qu'une fonctionnelle de $\overline{\mathbb{R}}^U$ est, en chaque point de U, plus grande que l'enveloppe supérieure de ses c-minorantes affines.

La propriété 3) est tout à fait similaire au théorème I.2 rencontré dans le cas convexe.

La propriété 4) traduit le fait que le plus grand minorant dans $\Gamma^c(U)$ d'une fonctionnelle de $\bar{\mathbb{R}}^U$ est l'enveloppe supérieure de ses c-minorantes affines.

Démonstration des propriétés

Propriété 1) : Cette inégalité découle immédiatement de la définition de la c-conjuguée d'une fonctionnelle a de $\bar{\mathbb{R}}^U$.

En effet, $a^c(v) \overset{déf}{=} \sup_{u \in U} (c(u,v) - a(u)) \quad \forall v \in V$

par conséquent,

$$a^c(v) \geq c(u,v) - a(u) \quad \forall u \in U, \forall v \in V.$$

Propriété 2) : soit a une fonctionnelle appartenant à $\bar{\mathbb{R}}^U$.

Par définition, $a^{cc}(u) = \sup_{v \in V} (c(u,v) - a^c(v)) \quad \forall u \in U$

En remplaçant $a^c(v)$ par sa valeur nous obtenons:

$$a^{cc}(u) = \sup_{v \in V} \left(c(u,v) - \sup_{u' \in U} (c(u',v) - a(u')) \right) \quad \forall u \in U$$

Ce qui implique :

$$a^{cc}(u) \leq \sup_{v \in V} (c(u,v) - c(u,v) + a(u)) \quad \forall u \in U$$

donc, $\quad a^{cc}(u) \leq a(u) \quad \forall u \in U$

Propriété 3) :

Condition nécessaire : $\forall a \in \bar{\mathbb{R}}^U$, si $a = a^{cc}$ alors $a \in \Gamma^c(U)$

Cette proposition est immédiate car,

$$\forall a \in \bar{\mathbb{R}}^U \quad a^{cc} \in \Gamma^c(U) \qquad \text{(cfr égalité (1) page 39)}$$

condition suffisante : $\forall a \in \overline{\mathbb{R}}^U$, si $a \in \Gamma^c(U)$ alors $a = a^{cc}$

Soit a une fonctionnelle quelconque de $\Gamma^c(U)$. Il existe donc un ensemble I, une famille $\{v_i\}_{i \in I}$ de points de V et une famille de nombres de $\mathbb{R}$ $\{\eta_{i \in I}\}$ tels que :

$$a(u) = \sup_{i \in I} (c(u, v_i) + \eta_i) \qquad \forall u \in U \qquad (1)$$

donc,

$$a^c(v) = \sup_{u \in U} \left(c(u, v) - \sup_{i \in I}(c(u, v_i) + \eta_i)\right) \qquad \forall v \in V$$

Fixons un élément arbitraire de I, soit j, nous avons :

$$a^c(v_j) = \sup_{u \in U} \left(c(u, v_j) - \sup_{i \in I}(c(u, v_i) + \eta_i)\right)$$

par conséquent,

$$a^c(v_j) \leq \sup_{u \in U} (c(u, v_j) - c(u, v_j) - \eta_j)$$

Comme j était arbitraire, nous en déduisons :

$$a^c(v_j) \leq -\eta_j \qquad \forall j \in I$$

or, l'inégalité de Young nous dit que :

$$a^{cc}(u) \geq c(u, v_j) - a^c(v_j) \qquad \forall j \in I, \forall u \in U$$

Il s'ensuit :

$$a^{cc}(u) \geq \sup_{i \in I} (c(u, v_i) + \eta_i) \qquad \forall u \in U$$

Par conséquent, en employant (1) et la propriété 2) nous pourrons conclure :

$$a^{cc}(u) = a(u) \qquad \forall u \in U$$

Propriété 4) : Soit a une fonctionnelle quelconque de $\bar{R}^U$ et b une fonctionnelle appartenant à $\Gamma^c(U)$ telle que :

$$a^{cc}(u) \leq b(u) \leq a(u) \qquad \forall u \in U$$

Comme b appartient à $\Gamma^c(U)$, $b = b^{cc}$

donc, $\quad b^{cc}(u) \geq a^{cc}(u) \qquad \forall u \in U \qquad (1)$

or, $\quad a(u) \geq b(u) \;\; \forall u \in U \implies a^{cc} \geq b^{cc}(u) \;\; \forall u \in U \;\; (2)$

car, $\quad a^c(v) \overset{\text{déf}}{=} \sup\limits_{u \in U} (c(u,v) - a(u)) \;\; \forall v \in V$

par hypothèse, $a(u) \geq b(u) \;\; \forall u \in U$

donc, $a^c(v) \leq \sup\limits_{u \in U} (c(u,v) - b(u)) \;\; \forall v \in V$

et par conséquent, $a^c(v) \leq b^c(v) \;\; \forall v \in V$

comme $a^{cc} \overset{\text{déf}}{=} (a^c)^c$, (2) est bien vérifiée.

suite aux inégalités (1) et (2) nous obtenons :

$$a^{cc} = b^{cc} = b$$

Nous voudrions maintenant obtenir une caractérisation plus précise de l'ensemble $\Gamma^c(U)$ similaire à celle donnée par le théorème I.1 pour le cas convexe.
Cela nous permettrait d'identifier facilement les fonctionnelles de $\overline{\mathbb{R}}^U$ qui coïncident avec l'enveloppe supérieure de leurs c-minorantes affines.

Pour obtenir un tel résultat, nous avons besoin de particulariser la fonctionnelle couplante c.
C'est à ce niveau que nous allons introduire le caractère pointu dont nous avions parlé dans les motivations du chapitre II.

<u>Définition II.2</u>

Supposons que l'espace U est muni d'une topologie. Une fonctionnelle couplante c des espaces U et V est dite POINTUE en un élément u_0 de U si :

$\forall N$ voisinage de u_0,

$\forall v \in V$

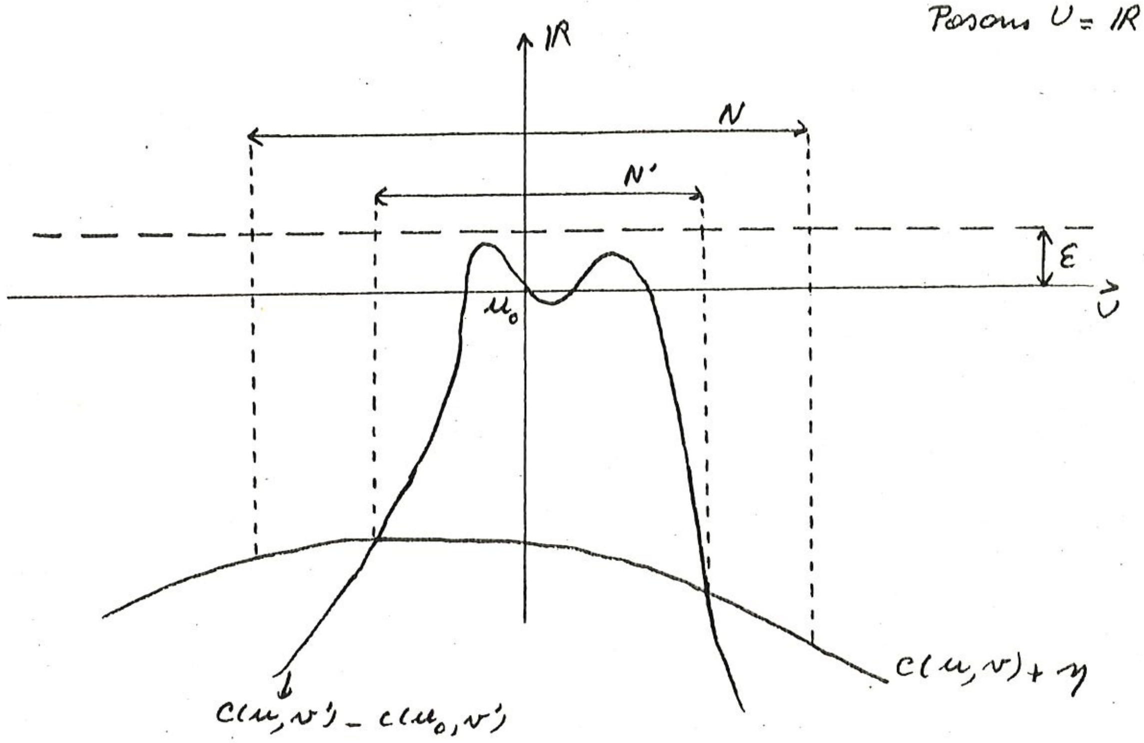
$\forall \eta, \varepsilon \in \mathbb{R}$, $\varepsilon > 0$

$\exists v' \in V$ et $\exists N'$ voisinage de u_0 tels que :

a) N' est inclus dans N, et

b)
$$\begin{cases} \forall u \notin N' & c(u,v') - c(u_0,v') \leq c(u,v) + \eta \\ \forall u \in N' & c(u,v') - c(u_0,v') \leq \varepsilon \end{cases}$$

Posons $U = \mathbb{R}$



A l'étude du dessin nous remarquons que c est pointue en u_0 si, considérée comme fonction de u, elle présente (pour certains éléments v de V) une pointe aussi fine que nous le voulons en u_0.
(Cela se voit en diminuant progressivement N, η et ε.)

Remarque 1:

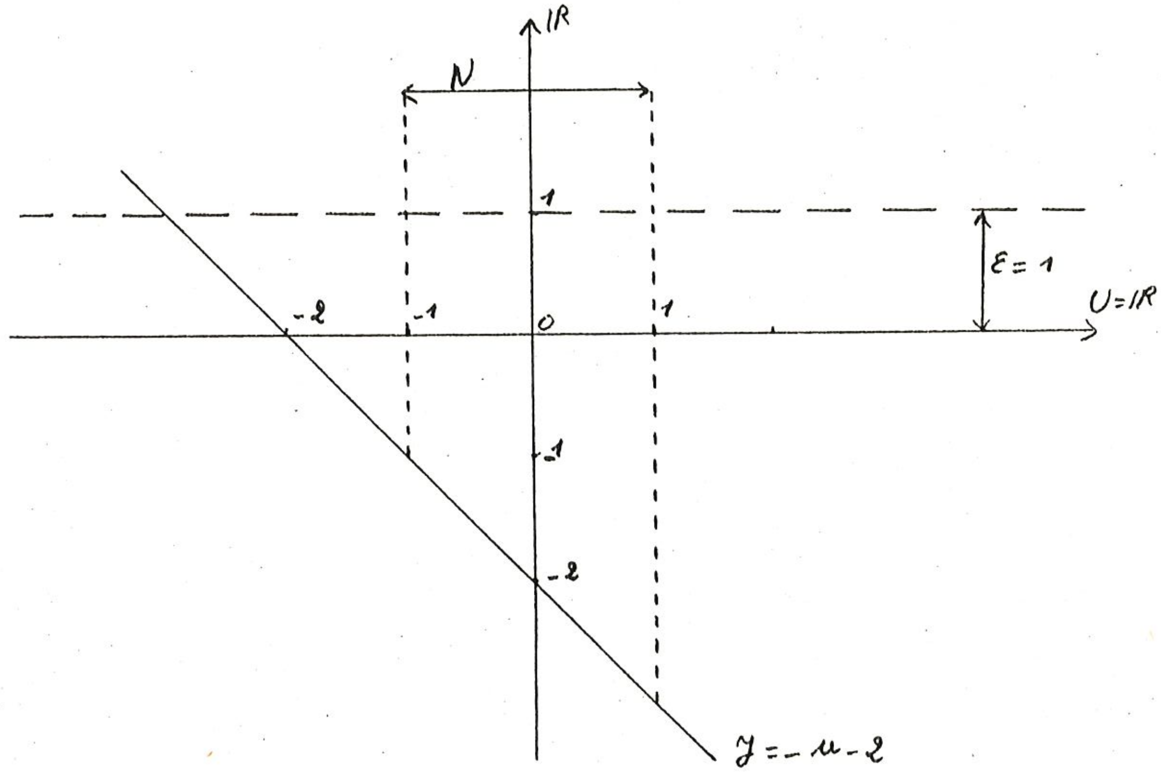
Considérons $U = V = \mathbb{R}$ et le produit scalaire comme fonctionnelle couplante. (nous reprenons ici un exemple rencontré dans le cas convexe.)

Le produit scalaire n'est pointu en aucun point de $\mathbb{R}$.

En effet, dans le cas particulier de l'origine de U, nous le

voyons immédiatement sur le dessin ci dessous.
(La représentation graphique serait tout à fait similaire pour tout autre point de U)

Soient $v = -1$, $N = [-1, 1]$, $\varepsilon = 1$, $y = -2$



Vérifier la définition revient dans notre cas à trouver une droite passant par l'origine qui, dans un voisinage N' inclus dans $[-1, 1]$, prend des valeurs inférieures à 1 et, à l'extérieur de N' minore la droite d'équation $y = -u - 2$. Nous voyons immédiatement que c'est impossible.

Remarque 2 :

Parmi les fonctionnelles pointues en u_0, nous avons celles dites de TYPE AIGUILLE en u_0 qui vérifient la définition II.3 en fixant $\varepsilon = 0$.

Nous possédons un lemme donnant des conditions suffisantes pour qu'une fonctionnelle de $\mathbb{R}^{U\times V}$ soit de type aiguille en u_0. (voir annexe I).

Nous allons donner deux exemples de fonctionnelles de type aiguille en un élément u_0 de U où nous appliquons ce lemme pour démontrer leur caractère pointu en u_0.

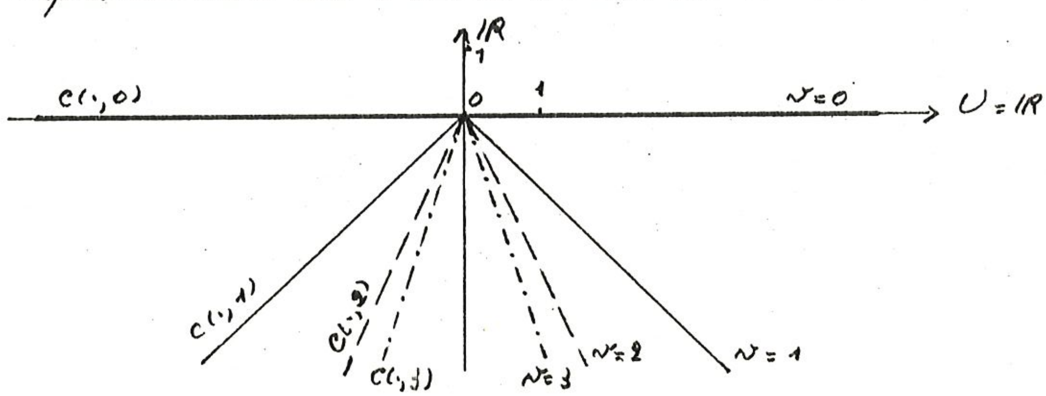
Exemple 1.a

Soient $U = \mathbb{R}^m$, $V = \mathbb{R}_+^m$

$$C(u,v) = -\sum_{i=1}^{m} v_i\,|u_i| \qquad u = (u_1, u_2 \ldots u_m) \in \mathbb{R}^m$$
$$v = (v_1, v_2, \ldots v_m) \in \mathbb{R}_+^m$$

C est de type aiguille à l'origine de $\mathbb{R}^m$ (démonstration voir annexe I.)

Représentation de C dans le cas où $m = 1$



Nous remarquons sur le dessin que, plus v croît, plus $c(.,v)$ devient pointue en 0.

Exemple 2.a

Soient $U = H$, un espace de Hilbert et $V = \mathbb{R}_+ \times H$

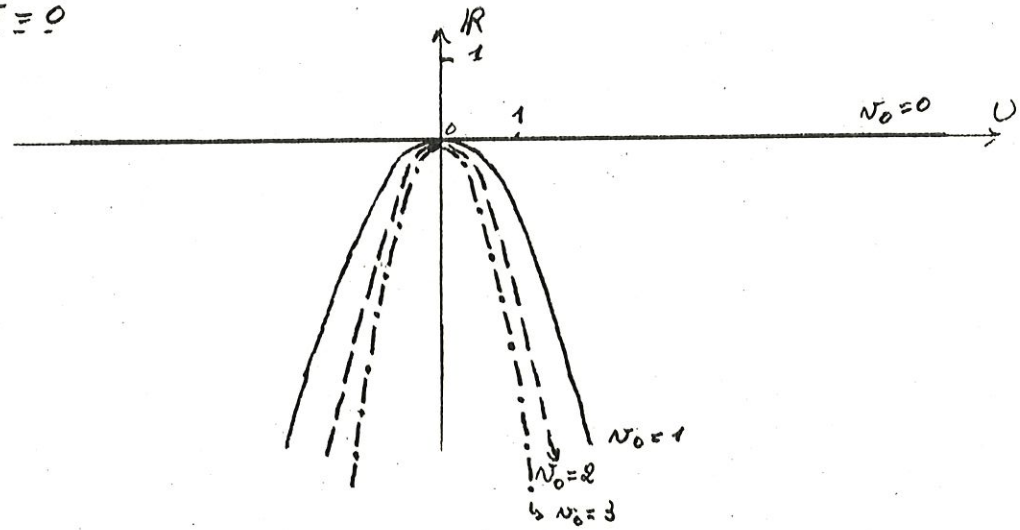
$$c(u,v) = -v_0 |u|^2 - \langle w, u \rangle \qquad v = (v_0, w) \in \mathbb{R}_+ \times H$$
$$u \in H$$

$\langle .,. \rangle$ désigne le produit scalaire défini dans H et $|.|$, la norme engendrée par le produit scalaire dans H.

c est de type aiguille en tout point de H
(démonstration voir annexe I.)

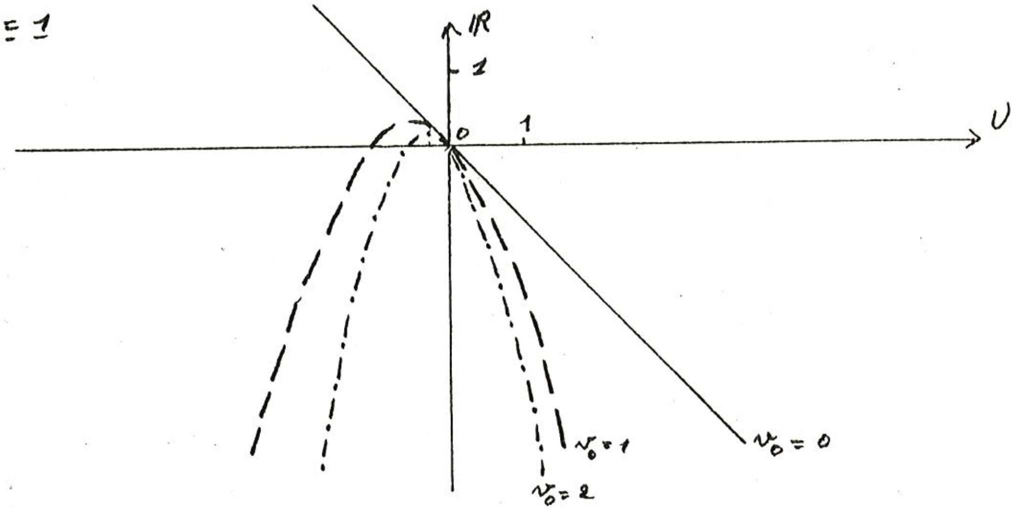
Représentation de c pour différentes valeurs de v_0 et de w dans le cas où $H = \mathbb{R}$.

1) $w = 0$



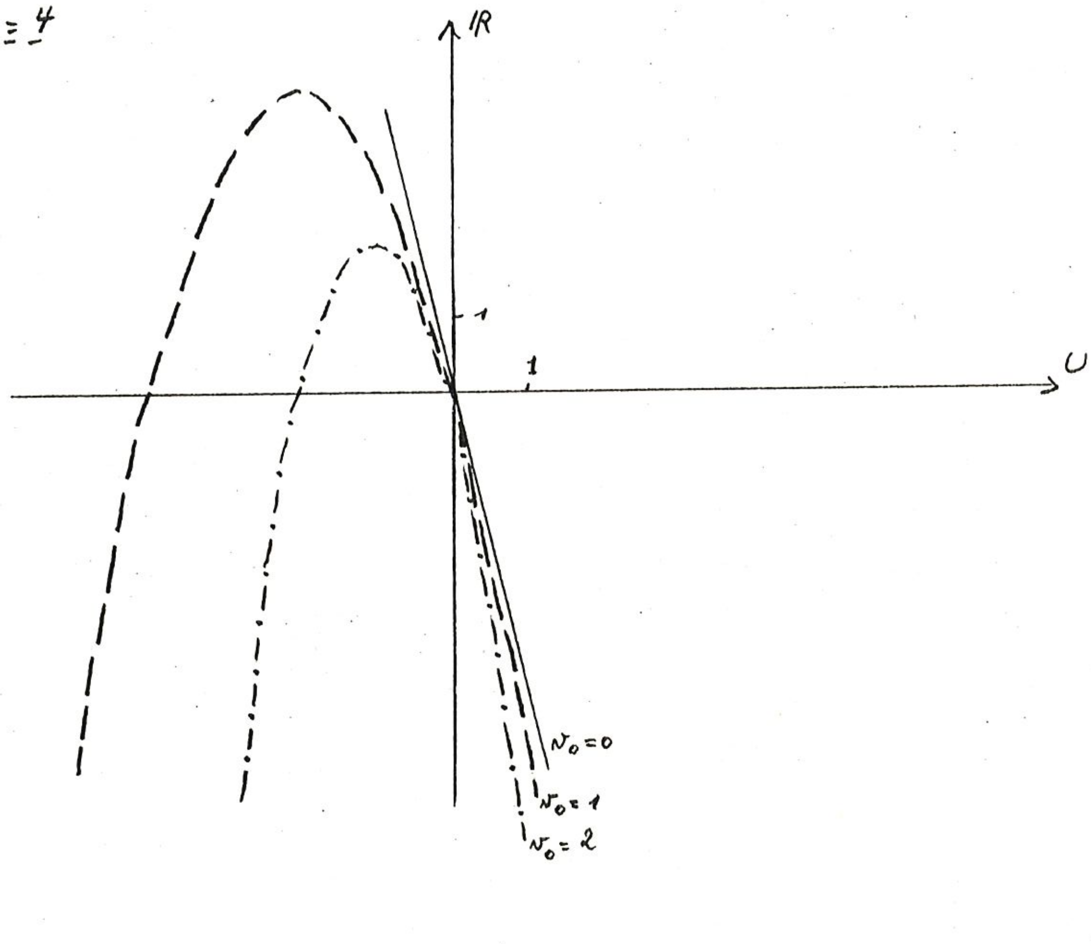
Pour $w = 0$, les graphes successifs $c(.,v)$ pour différentes valeurs de v_0 sont des aiguilles paraboliques qui s'affinent progressivement pour v_0 croissant.

2) $\underline{w = 1}$



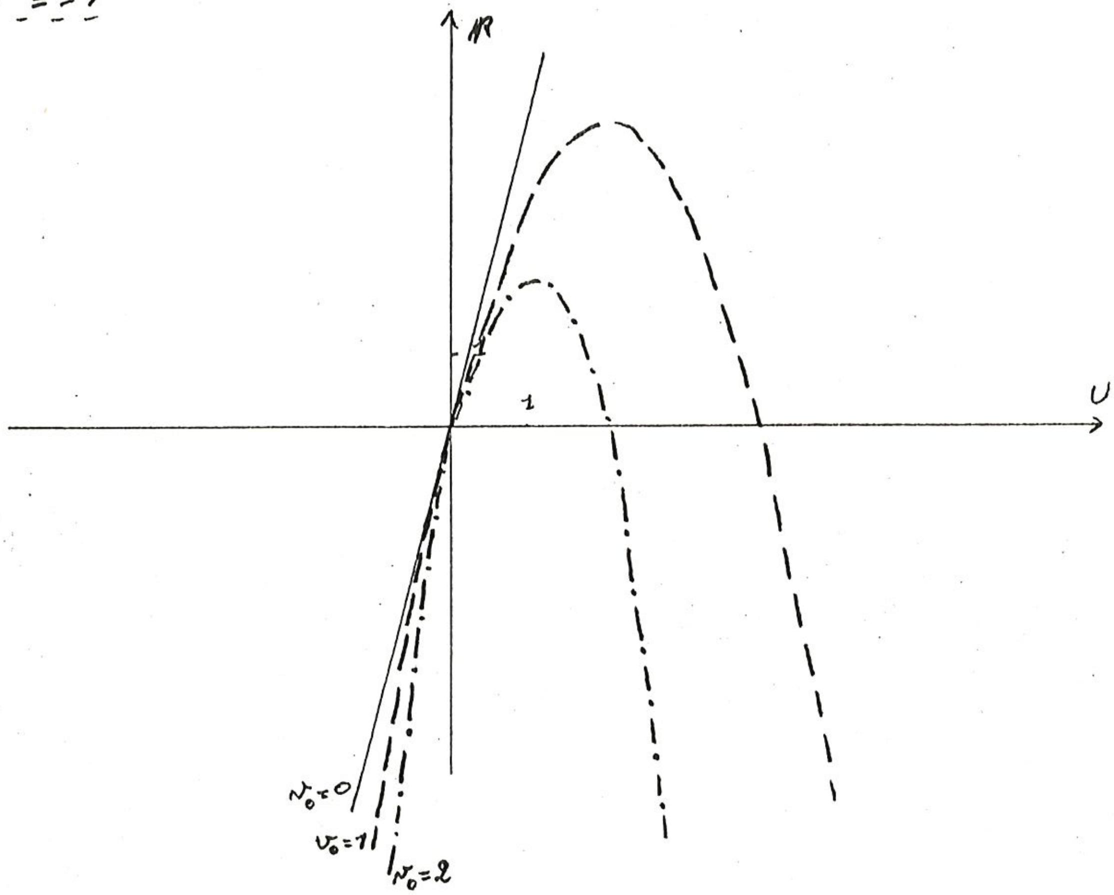
Pour $w = 1$, les aiguilles paraboliques sont légèrement décalées par rapport aux précédentes.

3) $\underline{w = 4}$



Pour w croissant nous remarquons que le décalage des aiguilles paraboliques s'accentue progressivement.

4) $w = -4$



Pour les valeurs négatives de w nous obtenons les mêmes courbes que pour les valeurs positives mais placées symétriquement par rapport à l'axe réel.

Nous allons maintenant introduire deux nouvelles notions avant de donner la caractérisation de $\Gamma^c(U)$

<u>Définition II.3</u>

Nous dirons qu'une fonctionnelle de $\overline{\mathbb{R}}^U$ est C-TEMPÉRÉE si et seulement si elle est minorée par une c-fonctionnelle élémentaire finie.

<u>Proposition II.1</u>

> Une fonctionnelle a de $\overline{\mathbb{R}}^U$ est c-tempérée si et seulement si il existe un élément u' de U tel que : $a^{cc}(u') \neq -\infty$

<u>Démonstration</u>

condition nécessaire

Soit a une fonctionnelle c-tempérée de $\overline{\mathbb{R}}^U$. Il existe donc un élément v de V et un réel y tels que :

$$c(u,v) + y \leq a(u) \qquad \forall u \in U$$

or

$$a^{cc}(u) = \sup_{\substack{v \in V \\ y \in \mathbb{R} \\ c(u,v)+y \leq a(u) \ \forall u \in U}} (c(u,v)+y) \qquad \forall u \in U$$

Par conséquent,

$$-\infty < c(u,v) + y \leq a^{cc}(u) \qquad \forall u \in U$$

condition suffisante

Soit a une fonctionnelle de $\overline{\mathbb{R}}^U$ telle qu'il existe un élément u' de U, vérifiant $a^{cc}(u') \neq -\infty$.

Par définition, $a^{cc}(u)$ est la valeur en u de l'enveloppe supérieure des c-minorantes affines de a.
Donc $a^{cc}(u') \neq -\infty$ implique qu'il existe une c-minorante affine de a (supposons-la définie par l'élément v de V et le réel η) telle que :

$$\begin{cases} -\infty < c(u', v) + \eta & (1) \\ c(u, v) + \eta \leq a(u) & \forall u \in U \end{cases}$$

Puisque c est à valeurs réelles, η est strictement supérieur à $-\infty$ (cfr inégalité (1)).
Par conséquent, a est c-tempérée.

<u>Définition II.4</u>

Supposons U muni d'une topologie.
Nous appellerons ENVELOPPE S.C.I (semi-continue inférieurement) d'une fonctionnelle a de $\overline{\mathbb{R}}^U$ la fonctionnelle notée $\bar{a}$ telle que :

$$\bar{a}(u) \overset{\text{déf}}{=} \liminf_{u' \to u} a(u') = \sup_{V \in \mathcal{V}(u)} \inf_{u' \in V} a(u') \quad \forall u \in U$$

où $\mathcal{V}(u)$ désigne la famille des voisinages de u.

Remarquons que $\bar{a}$, de par sa définition, est une fonctionnelle s.c.i sur U.

Proposition II.2

> Soit une fonctionnelle a de $\overline{\mathbb{R}}^U$.
> $\bar{a}$ est la plus grande minorante s.c.i de a.

Démonstration

Soit une fonctionnelle b de $\overline{\mathbb{R}}^U$ qui soit une minorante s.c.i de a.

Donc,

$$b(u) \leq a(u) \qquad \forall u \in U$$

Par conséquent, pour un élément u' arbitraire dans U nous avons :

$$\operatorname*{Inf}_{u \in V} b(u) \leq \operatorname*{Inf}_{u \in V} a(u) \qquad \forall V \in \mathcal{F}(u')$$

($\mathcal{F}(u')$ désignant la famille des voisinages de u',

Il s'ensuit :

$$\bar{b}(u') \leq \bar{a}(u')$$

En outre, puisque b est s.c.i, $b = \bar{b}$

Nous en déduisons :

$$b(u') \leq \bar{a}(u') \qquad (1)$$

u' étant arbitraire, l'inégalité (1) est vérifiée pour tout u dans U et la théo est démontrée.

Toutes les notions introduites précédemment vont maintenant nous permettre de donner quelques relations d'égalité entre une fonctionnelle de $\bar{\mathbb{R}}^U$ et l'enveloppe supérieure de ses c-minorantes affines ainsi que la caractérisation de $\Gamma^c(U)$.

Théorème II.1

Si c est une fonctionnelle de $\mathbb{R}^{U\times V}$ pointue en u_0, un élément de U, alors toute fonctionnelle c-tempérée a de $\bar{\mathbb{R}}^U$, s.c.i en u_0, vérifie l'égalité suivante :

$$a(u_0) = a^{cc}(u_0)$$

En outre, si pour tout v appartenant à V, $c(.,v)$ est s.c.i en u_0 alors pour toute fonctionnelle c-tempérée a de $\bar{\mathbb{R}}^U$ nous avons :

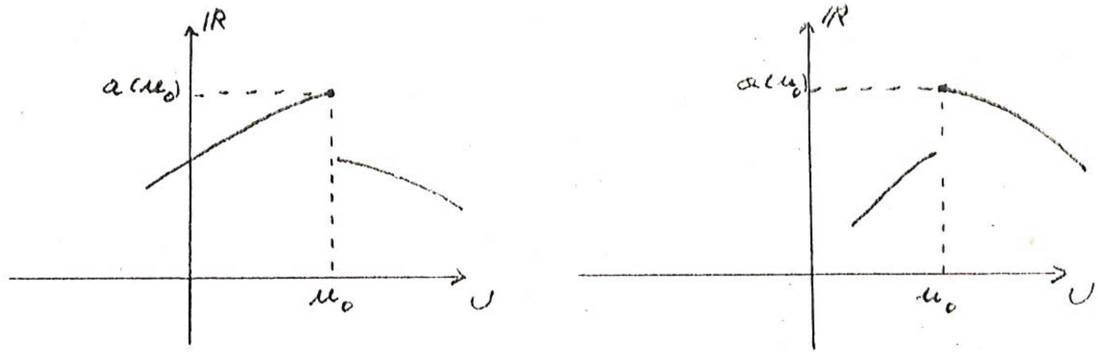
$$\bar{a}(u_0) = a^{cc}(u_0)$$

(démonstration voir annexe II.)

Nous allons essayer d'interpréter géométriquement les hypothèses de ce théorème et de sentir par ce biais leur utilité.

. a c-tempérée implique $a^{cc}(u_0) \neq -\infty$ (cfr condition nécessaire de la proposition II.1)
Dans le cas contraire, à moins que $a(u_0)$ ne soit $-\infty$, nous n'aurons pas l'égalité entre $a(u_0)$ et $a^{cc}(u_0)$.

La semi-continuité inférieure de a en u_0 empêche un certain type de discontinuité qui pourrait rendre $a(u_0)$ difficilement accessible par le biais des c-minorantes affines.
Par exemple, dans le cas où $U = \mathbb{R}$, cela empêche les discontinuités représentées ci-dessous.



Si on évite des discontinuités de ce genre, grâce au caractère pointu de c en u_0 nous pourrons atteindre $a(u_0)$ en prenant la borne supérieure en u_0 des c-minorantes affines de a.

Dans le second théorème, nous allons donner un résultat analogue au théorème I.1 rencontré dans le cas convexe.

Théorème II.2

Si c est une fonctionnelle de $\mathbb{R}^{U\times V}$ pointue en chaque point de U alors, toute fonctionnelle de $\bar{\mathbb{R}}^U$ s.c.i et c-tempérée appartient à $\Gamma^c(U)$
En outre, si pour tout v appartenant à V, $c(\cdot, v)$ est s.c.i sur U alors toute fonctionnelle a de $\bar{\mathbb{R}}^U$ ne prenant jamais la valeur $(-\infty)$ satisfait l'équivalence suivante :

$$a \in \Gamma^c(U) \quad \text{ssi} \quad a \text{ est s.c.i et c-tempérée}$$

La démonstration de ce théorème se trouve dans l'annexe II et est essentiellement basée sur le théorème II.1 les interprétations géométriques des hypothèses du théorème II.1 restent donc tout à fait valables ici.

Nous arrivons maintenant à un résultat essentiel qui va nous montrer que la définition d'une fonctionnelle pointue en un point est bien choisie.
le théorème II.2 nous disait que, si c est pointue et s.c.i sur V alors $\Gamma^c(V)$ est identique à l'ensemble des fonctionnelles de $\overline{\mathbb{R}}^V$ s.c.i et c-tempérées.
Dans le théorème II.3 nous aurons en quelque sorte une réciproque de ce théorème. C'est à dire, si nous voulons que $\Gamma^c(V)$ coïncide avec l'ensemble des fonctionnelles de $\overline{\mathbb{R}}^V$ s.c.i et c-tempérées, il est presque obligatoire que c soit pointue sur V

Théorème II.3

Si c est une fonctionnelle de $\mathbb{R}^{V\times V}$ telle que pour tout v appartenant à V $c(\cdot,v)$ est s.c.s en u_0, un élément de V, et pour toute fonctionnelle a de $\overline{\mathbb{R}}^V$ s.c.i en u_0, c-tempérée nous avons : $a(u_0) = a^{cc}(u_0)$
alors c est pointue en u_0.

(démonstration voir annexe II.)

Si nous résumons les résultats des théorèmes II.2 et II.3, nous obtenons le corollaire suivant :

Corollaire

Si c est une fonctionnelle de $\mathbb{R}^{U \times V}$ telle que, pour tout v appartenant à V, $c(\cdot, v)$ est continue sur U alors les deux propositions suivantes sont équivalentes :

1) $\forall a \in \overline{\mathbb{R}}^U$ t.q $a(u) \neq -\infty \quad \forall u \in U$

$a \in \Gamma^c(U)$ ssi a est s.c.i et c-tempérée

2) c est pointue en chaque point de U.

Nous allons maintenant introduire une dernière notion avant de définir la dualité en optimisation non convexe.

Définition II.5

Soit ε, un réel strictement positif.

Une fonctionnelle a de $\overline{\mathbb{R}}^U$ est dite ε-c-SOUS-DIFFÉRENTIABLE en un élément u_0 de U si $a(u_0)$ est fini et si il existe v_0 appartenant à V tel que :

$$a(u) - a(u_0) \geq c(u, v_0) - c(u_0, v_0) - \varepsilon \qquad \forall u \in U$$

Si nous exprimons cette définition en termes géométriques nous obtenons l'énoncé suivant :

Une fonctionnelle a de $\overline{\mathbb{R}}^U$ est ε-c-sous différentiable en un élément u_0 de U si il existe une c-fonctionnelle élémentaire finie qui minore a et dont la valeur en u_0 diffère de ε par rapport à $a(u_0)$.

Démonstration de l'équivalence des deux énoncés

condition nécessaire :

Soit a une fonctionnelle de $\overline{\mathbb{R}}^U$ finie en un élément u_0 de U et telle qu'il existe v_0 dans V satisfaisant l'inégalité suivante pour un réel ε fixé :

$$a(u) - a(u_0) \geqslant c(u, v_0) - c(u_0, v_0) - \varepsilon \qquad \forall u \in U$$

Posons $\eta = a(u_0) - c(u_0, v_0) - \varepsilon$.

η est fini car $a(u_0)$, ε et $c(u_0, v_0)$ appartiennent à $\mathbb{R}$.

Donc, $c(u, v_0) + \eta$ est une c-fonctionnelle élémentaire finie qui minore a et qui prend en u_0 la valeur $a(u_0) - \varepsilon$

condition suffisante :

Soit a une fonctionnelle de $\overline{\mathbb{R}}^U$ et supposons que $c(\cdot, v_0) + \eta$ est une c-fonctionnelle élémentaire minorant a et telle que :

$$c(u_0, v_0) + \eta = a(u_0) - \varepsilon$$

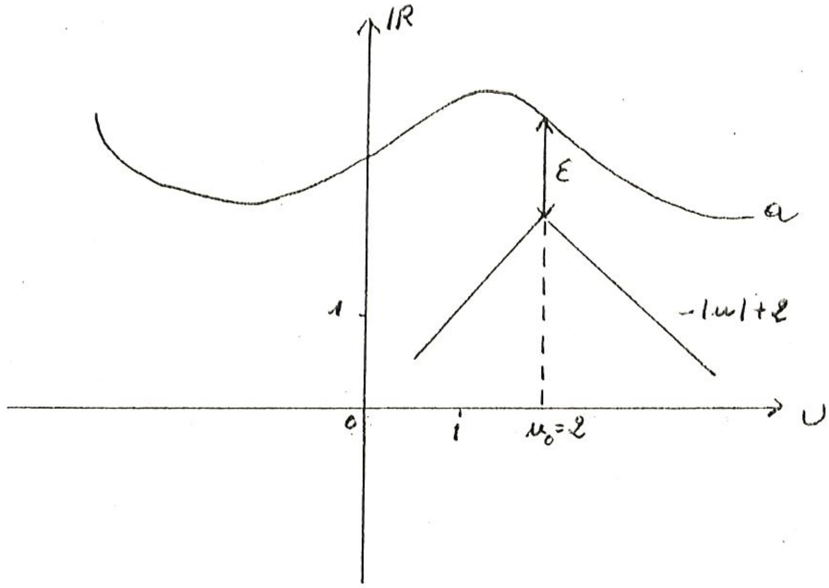
Comme $c(u_0, v_0)$, η, ε sont réels, $a(u_0)$ est fini.

En outre $\eta = a(u_0) - c(u_0, v_0) - \varepsilon$

donc, $\quad a(u) - a(u_0) \geqslant c(u, v_0) - c(u_0, v_0) - \varepsilon \quad \forall u \in U$

<u>Exemple</u>

Posons $U = V = \mathbb{R}$ et $c(u,v) = -v|u|$ $\quad \forall u \in U, \forall v \in V$



a est ε-c sous différentiable en $u_0 = 2$ car il existe un élément v_0 de V et un réel η (ici $v_0 = 1$ et $\eta = 2$ conviennent) tels que :

$$\left| \begin{array}{ll} c(u, v_0) + \eta \leq a(u) & \forall u \in U \\ c(u_0, v_0) + \eta = a(u_0) - \varepsilon & \end{array} \right. \qquad (2)$$

Un point v_0 de V vérifiant (2) est appelé ε-c-SOUS-GRADIENT de a en u_0.

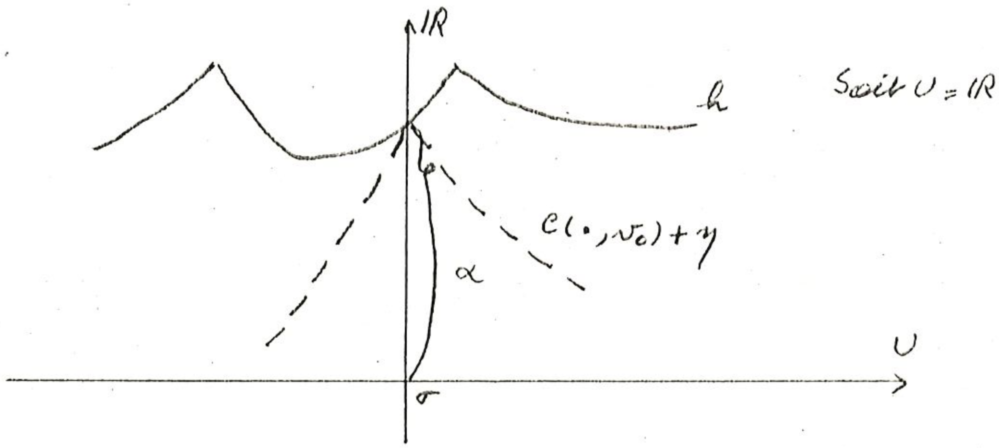
L'ensemble de tous les ε-c-sous-gradients de a en u_0 est appelé l'ε-c-SOUS-DIFFÉRENTIEL de a en u_0 et noté $c\text{-}\partial_\varepsilon a(u_0)$.

Dans le cas où ε est nul, nous omettons le préfixe et l'indice. Remarquons que les notions de c-sous-gradient et de c-sous-différentiel ne sont que des généralisations des notions de sous-gradient et de sous-différentiel définies dans le chapitre I. Il en sera de même pour les propriétés.

Remarque:

L'intérêt de ce concept réside dans le fait que la c-sous-différentiabilité d'une fonctionnelle a de $\bar{\mathbb{R}}^{U}$ en un élément u_0 de U implique automatiquement que, en u_0, cette fonctionnelle coïncide avec l'enveloppe supérieure de ses c-minorantes affines.

Par exemple, considérons le cas particulier de la fonction de perturbation h en σ (cfr II.1 motivations)
Si h est c-sous-différentiable en σ, α $(= h(\sigma))$ sera atteint en σ par une c-minorante affine de h.



$c(\cdot, v_0) + \eta$ représente la c-fonctionnelle élémentaire finie qui minore h et dont la valeur en σ est $h(\sigma)$ (cfr $\varepsilon = 0$)

Propriétés des ε-c-sous-différentiels

Soient a, une fonctionnelle de $\bar{\mathbb{R}}^U$, et ε, un réel positif.
Considérons u_0 un élément de U et v_0 appartenant à V.

1) $v_0 \in c\text{-}\partial_\varepsilon a(u_0)$ ssi $c(u_0, v_0) \leq a(u_0) + a^c(v_0) \leq c(u_0, v_0) + \varepsilon$

2) Si $c\text{-}\partial_\varepsilon a(u_0) \neq \emptyset$ alors $a^{cc}(u_0) \leq a(u_0) \leq a^{cc}(u_0) + \varepsilon$

3) Si $a(u_0) = a^{cc}(u_0)$ alors $c\text{-}\partial a(u_0) = c\text{-}\partial a^{cc}(u_0)$

4) les propositions suivantes sont équivalentes pour a appartenant à $\Gamma^c(U)$.

ⓐ $v_0 \in c\text{-}\partial a(u_0)$

ⓑ $a(u_0) + a^c(v_0) = c(u_0, v_0)$

ⓒ $u_0 \in c\text{-}\partial a^c(v_0)$

Commentaire :

La propriété 2) exprime que si il existe une c-fonctionnelle élémentaire finie qui minore a et dont la valeur en u_0 diffère de ε par rapport à $a(u_0)$ alors, en u_0, nous pouvons approcher a, à ε près, par l'enveloppe supérieure de ses c-minorantes affines

La propriété 3) nous dit que si a et l'enveloppe supérieure de ses c-minorantes affines (a^{cc}) coïncident en un point u_0 de U alors toute c-minorante affine de a exacte en u_0 est une c-minorante affine de a^{cc} exacte en u_0.

Démonstration des propriétés

Propriété 1 :

Par définition, $v_0 \in c\text{-}\partial_\varepsilon a(u_0)$ ssi

$$\begin{cases} a(u_0) \text{ est fini} \\ a(u) - a(u_0) \geq c(u, v_0) - c(u_0, v_0) - \varepsilon \quad \forall u \in U \quad (1) \end{cases}$$

or, l'inégalité (1) est équivalente à :

$$c(u_0, v_0) + \varepsilon \geq a(u_0) + (c(u, v_0) - a(u)) \quad \forall u \in U$$

Puisque c'est valable pour tout u dans U, nous avons.

$$c(u_0, v_0) + \varepsilon \geq a(u_0) + \sup_{u \in U} (c(u, v_0) - a(u))$$

Par conséquent, en utilisant la définition de $a^c(v_0)$ et l'inégalité de Young nous obtenons :

$$v_0 \in c\text{-}\partial_\varepsilon a(u_0) \quad \text{ssi} \quad c(u_0, v_0) \leq a(u_0) + a^c(v_0) \leq c(u_0, v_0) + \varepsilon$$

Propriété 2 :

Par hypothèse, il existe un élément v_0 de V tel que v_0 est un ε-c-sous-gradient de a en u_0.

Donc, par la propriété 1, nous déduisons l'inégalité suivante :

$$c(u_0, v_0) - a^c(v_0) \leq a(u_0) \leq c(u_0, v_0) - a^c(v_0) + \varepsilon$$

Par conséquent, en utilisant la définition de a^{cc} et la propriété 2 des fonctionnelles c-conjuguées nous obtenons :

$$a^{cc}(u_0) \leq a(u_0) \leq a^{cc}(u_0) + \varepsilon$$

Propriété 3 :

Soit v_0 appartenant à $c\text{-}\partial a(u_0)$. v_0 définit donc une c-minorante affine de a exacte en u_0.

Par la propriété 4 des fonctionnelles c-conjuguées, a^{cc} admet les mêmes c-minorantes affines que a.

Comme $a(u_0) = a^{cc}(u_0)$, v_0 définit une c-minorante affine de a^{cc} exacte en u_0 et par conséquent, v_0 appartient à $c\text{-}\partial a^{cc}(u_0)$
on procède de même pour la réciproque.

Propriété 4

Par la propriété 1, nous avons :

$v_0 \in c\text{-}\partial a(u_0)$ ssi $c(u_0, v_0) = a(u_0) + a^c(v_0)$

et

$u_0 \in c\text{-}\partial a^c(v_0)$ ssi $c(u_0, v_0) = a^c(v_0) + a^{cc}(u_0)$

Or, par hypothèse a appartient à $\Gamma^c(U)$ et donc $a = a^{cc}$
(cfr propriété 3 des fonctionnelles c-conjuguées)
Par conséquent, l'équivalence entre les trois propositions est démontrée.

Nous possédons maintenant tous les éléments théoriques nécessaires pour introduire la théorie de la dualité en optimisation non convexe.

II.3. Stabilité et dualité en optimisation non convexe

Dans ce paragraphe, nous allons appliquer les résultats obtenus précédemment dans le cadre des problèmes d'optimisation non convexes.

Nous considérons de nouveau le problème de minimisation (P) introduit dans le paragraphe II.1.

Remarque :

Durant toute la suite du travail, nous supposerons que la fonctionnelle couplante de U et V choisie est normalisée en σ c'est à dire :

$$c(\sigma, v) = 0 \qquad \forall v \in V$$

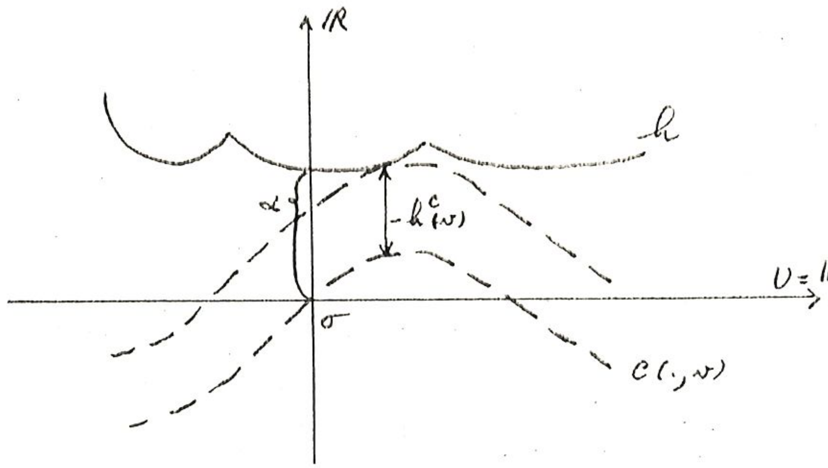
Posons $g = h^c(v)$ appartenant à $\Gamma^c(V)$

Nous définissons le problème de minimisation DUAL correspondant à la famille des problèmes de minimisation perturbés $\{(P_u) \mid u \in U\}$ par :

$$(Q) \qquad \inf_{v \in V} g(v) \overset{\text{noté}}{=} \beta$$

Remarques :

1) Similairement au cas convexe, le problème dual consiste, au signe près, à chercher la valeur maximale prise en σ par une c-minorante affine de h.
(Cette assertion découle de l'interprétation géométrique de h^c et de la normalisation de c en σ.)



2) Il est intéressant de remarquer que le problème dual peut être un problème de minimisation convexe.

Par exemple, ce sera le cas si V est un espace vectoriel et $c(u,.)$ est convexe sur V pour tout u dans U.

En effet, soient v_1, v_2 des éléments de V et λ appartenant à $]0,1[$, nous devons prouver l'inégalité suivante :

$$g(\lambda v_1 + (1-\lambda) v_2) \leq \lambda g(v_1) + (1-\lambda) g(v_2).$$

Comme V est un espace vectoriel, $\lambda v_1 + (1-\lambda) v_2 \in V$

En outre,

$$g(\lambda v_1 + (1-\lambda) v_2) \overset{\text{déf}}{=} \sup_{u \in U} [c(u, \lambda v_1 + (1-\lambda) v_2) - h(u)]$$

Donc, puisque $c(u,.)$ est convexe sur V, nous obtenons :

$$g(\lambda v_1 + (1-\lambda) v_2) \leq \sup_{u \in U} [\lambda c(u, v_1) + (1-\lambda) c(u, v_2) - \lambda h(u) - (1-\lambda) h(u)]$$

De cette inégalité nous déduisons la thèse.

Puisque h^{cc} est par définition l'enveloppe supérieure des c-minorantes affines de h, la remarque 1 implique $\beta = -h^{cc}(\sigma)$ et par conséquent,

(II.1) $$\boxed{-\alpha = -h(\sigma) \leq -h^{cc}(\sigma) = \beta}$$

Comme dans le cas convexe, si h appartient à $\Gamma^c(U)$ les valeurs optimales du problème dual et primal coïncident au signe près. (cfr propriété 3 des fonctionnelles c-conjuguées)

En outre, en utilisant la formule (II.1) et l'hypothèse que f est non identiquement égale à $(+\infty)$, nous avons :

(II.2) $$\boxed{\alpha < +\infty \text{ et } \beta > -\infty}$$

Grâce aux résultats (II. 1) et (II. 2) nous pouvons distinguer deux cas.

a) $\beta = +\infty$

Dans ce cas, h n'est pas c-tempérée (cfr proposition II. 1) elle ne possède donc pas de c-minorante affine.

Nous obtenons alors, si $\alpha \neq -\infty$, un écart infini entre α et $-\beta$.

b) β est fini

Ceci est équivalent à $h^{cc}(\sigma)$ fini. Donc, h admet une c-minorante affine (cfr proposition II. 1) et, suite au théorème II. 1 et à la formule (II. 1) nous obtenons le résultat suivant :

Théorème II. 4

Si $\beta < +\infty$ et si c est pointue en σ alors

h est s.c.i en σ entraine $-\alpha = \beta$

En outre, si $c(\cdot, v)$ est s.c.i en σ pour tout élément v de V

alors $$-\beta = \bar{h}(\sigma)$$

Reprenons la définition de "inf - stabilité" donnée dans le cas convexe. Nous obtenons alors l'équivalence suivante :

Théorème II.5

Si c est pointue en σ, $c(\cdot, v)$ s.c.i en σ pour tout v appartenant à V et $\beta < +\infty$ alors :

(P) est inf - stable si et seulement si $-\beta = \alpha$

(démonstration voir annexe II)

Nous obtenons presque l'équivalence donnée dans le théorème I.3 du cas convexe. La seule différence est l'hypothèse supplémentaire sur β qui nous assure que h admet des c - minorantes affines.

En outre, vous pouvez remarquer que nous ne parlons pas de l'équivalence entre l'inf stabilité de (P) et l'inf stabilité de (Q).

Cette équivalence dans le cas convexe découlait de la symétrie du problème.

Cette théorie de dualité en optimisation non convexe n'est pas symétrique. Il existe, cependant, des théories de la dualité en optimisation non convexe qui le sont c'est à dire où le dual du problème dual est le primal.

(cfr [4])

Pour pouvoir parler du dual du dual et donc d'inf-stabilité de (Q) dans ce cadre, nous devrions introduire un nouvel espace Y et une autre fonctionnelle couplante définie sur $X \times Y$.

Nous n'envisagerons pas cet aspect du problème dans ce travail.

En appliquant les propriétés des fonctions c-conjuguées et des c-sous-différentiels dans le cadre de la dualité, nous obtenons quelques propositions intéressantes.

Proposition II.3

Si $\beta < +\infty$, l'ensemble des solutions optimales du dual noté B coïncide avec le c-sous-différentiel de h^{cc} en σ.

Démonstration

Par définition, $B = \{\bar{v} \in V \mid g(\bar{v}) = \beta\}$

Considérons un élément $\bar{v}$ appartenant à $c\text{-}\partial h^{cc}(\sigma)$.

Par la propriété 1 des c-sous-différentiels ceci est équivalent à :

$$h^{cc}(\sigma) + h^{ccc}(\bar{v}) = c(\sigma, \bar{v})$$

Comme c est normalisée en σ et h^c appartient à $\Gamma^c(U)$ cela revient à :

$$h^{cc}(\sigma) + h^c(\bar{v}) = 0$$

Or, $\beta = -h^{cc}(\sigma)$ et $g = h^c$.

Donc, $\bar{v} \in c\text{-}\partial h^{cc}(\sigma)$ ssi $g(\bar{v}) = \beta$

ou, $\bar{v} \in c\text{-}\partial h^{cc}(\sigma)$ ssi $\bar{v} \in B$.

Ce résultat se conçoit aisément du point de vue géométrique grâce à l'interprétation du dual et de $c\text{-}\partial h^{cc}(\sigma)$. Si $\bar{v}$ appartient à B, on peut lui associer une c-minorante affine de h dont la valeur en σ est $(-\beta)$. h^{cc} étant l'enveloppe supérieure des c-minorantes affines de h, il est évident, puisque $h^{cc}(\sigma) = -\beta$, que $\bar{v}$ est un élément de $c\text{-}\partial h^{cc}(\sigma)$.

<u>Proposition II.4</u>

> La c-sous-différentialité de h en σ est équivalente à l'existence d'une solution duale et à l'égalité entre $-\beta$ et α finis dans ce cas ; c'est à dire :
>
> $$c\text{-}\partial h(\sigma) \neq \emptyset \text{ ssi } \mathcal{B} \neq \emptyset \text{ et } -\alpha = \beta < +\infty$$

<u>Démonstration</u>

<u>condition nécessaire</u>

Supposons que $c\text{-}\partial h(\sigma)$ est non vide.

Par la propriété 2 des c-sous-différentiels cela implique :

$$h^{cc}(\sigma) = h(\sigma) \text{ et } h(\sigma) \text{ fini}$$

Donc, en utilisant les formules (II.1) et (II.2) (page 65) nous avons :

$$-\alpha = \beta < +\infty.$$

Comme $h^{cc}(\sigma) = h(\sigma)$ nous pouvons utiliser la propriété 3 des c-sous-différentiels et par conséquent,

$$c\text{-}\partial h(\sigma) = c\text{-}\partial h^{cc}(\sigma) = \mathcal{B}. \quad \text{(cfr proposition II.3)}$$

<u>condition suffisante</u>

Supposons que $\mathcal{B}$ est non vide et $-\alpha = \beta < +\infty$.

Donc, $h(\sigma) = h^{cc}(\sigma)$ (cfr formule II.1 page 65)

En utilisant de nouveau la propriété 3 des c-sous-différentiels nous obtenons :

$$c\text{-}\partial h(\sigma) = c\text{-}\partial h^{cc}(\sigma) = \mathcal{B} \quad \text{(cfr proposition II.3)}$$

De nouveau, nous pouvons interpréter géométriquement le résultat précédent.

$-\alpha = \beta < \infty$ signifie que, en σ, h coïncide avec la borne supérieure de ses c-minorantes affines.

Si, en outre, B est non vide cette borne est atteinte c'est à dire, il existe une c-minorante affine de h dont la valeur en σ est $\alpha = h(\sigma)$, donc $c\text{-}\partial h(\sigma)$ est non vide.

Nous allons maintenant introduire le concept du LAGRANGIEN l de (P).

Nous le définissons par :

$$l(x,v) \overset{\text{déf}}{=} \sup_{u \in U} \left[c(u,v) - \phi(x,u) \right] \qquad x \in X,\ v \in V$$

Remarquons que $l_x = \phi_x^c$ et possède par conséquent la même interprétation géométrique.

(l_x et ϕ_x désignent les fonctions partielles déduites respectivement de l et de ϕ pour un élément x fixé dans X)

<u>Propriétés du lagrangien</u>

1) $\beta = \inf_{v \in V} \sup_{x \in X} l(x,v)$

2) $-f(x) \leq \inf_{v \in V} l(x,v) \qquad \forall x \in X$

Remarquons que la propriété 2 entraine : $-\alpha \leq \sup_{x \in X} \inf_{v \in V} l(x,v)$

mais, contrairement au cas convexe nous n'avons pas l'égalité.

Démonstration des propriétés :

Propriété 1

$$\sup_{x \in X} l(x,v) = \sup_{x \in X} \sup_{u \in U} [c(u,v) - \phi(x,u)]$$

$$= \sup_{u \in U} \sup_{x \in X} [c(u,v) - \phi(x,u)]$$

$$= \sup_{u \in U} [c(u,v) - \inf_{x \in X} \phi(x,u)]$$

$$= \sup_{u \in U} [c(u,v) - h(u)]$$

$$= h^c(v) = g(v)$$

Par conséquent, suite à la définition de β, nous obtenons :

$$\beta = \inf_{v \in V} \sup_{x \in X} l(x,v)$$

Propriété 2

Par définition de ϕ nous avons :

$$-f(x) = -\phi(x,\sigma) \qquad \forall x \in X$$

Il s'ensuit, vu la normalisation de c en σ, que :

$$-f(x) = c(\sigma,v) - \phi(x,\sigma) \qquad \forall x \in X,\ \forall v \in V.$$

donc, par définition de l,

$$-f(x) \leq l(x,v) \qquad \forall x \in X,\ \forall v \in V$$

et par conséquent,

$$-f(x) \leq \inf_{v \in V} l(x,v)$$

Voici maintenant deux théorèmes importants donnant des conditions suffisantes pour qu'un élément $\bar{v}$ de V ($\bar{x}$ de X) soit solution du dual (du primal).

Théorème II.6

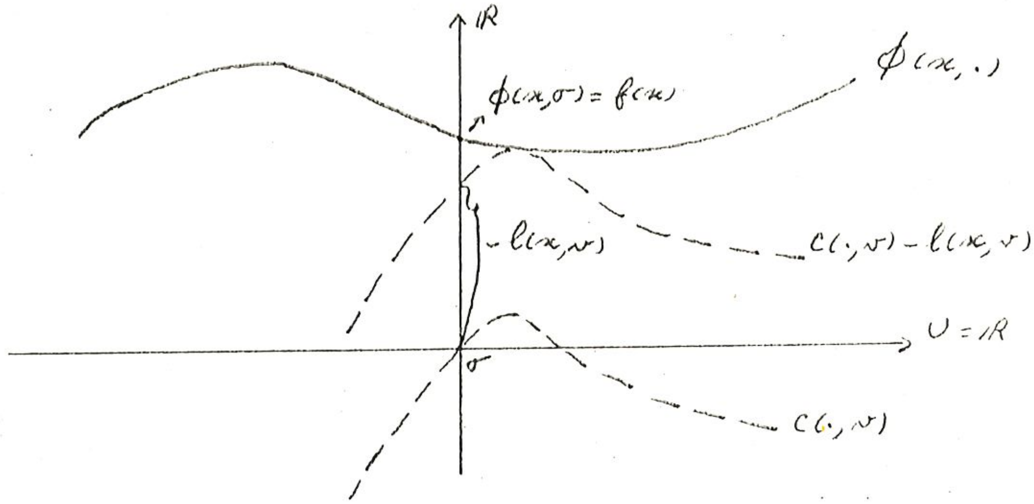
Si $(\bar{x}, \bar{v})$ est un point de selle du lagrangien ($\bar{x} \in X$, $\bar{v} \in V$) alors $\bar{v}$ est une solution du problème dual (Q).
Dans le cas où, $-\alpha = \beta < +\infty$, c pointue en σ et $\phi(\bar{x}, .)$ s.c.i en σ, nous avons aussi que $\bar{x}$ est une solution du problème primal (P).
Si, en outre, $c(., v)$ est s.c.i en σ pour tout v dans V alors :

$\bar{x}$ est une solution de (P) ssi $\phi(\bar{x}, .)$ est s.c.i en σ.

(démonstration voir annexe III)

Nous retrouvons ici la correspondance entre les points de selle du lagrangien et les couples de $X \times Y$ dont la première composante est une solution de (P) et la seconde, une solution de (Q). (nous l'avions déjà rencontrée dans le cas convexe cfr théorème I.4)

Essayons de justifier géométriquement ce résultat.
Rappelons-nous d'abord que $l(x, v) = \phi_x^c(v) \quad \forall (x, v) \in X \times V$
Donc, $l(x, v)$ donne au signe près la plus grande valeur prise en σ par une c-minorante affine de $\phi_x(.)$ définie par v.



$(\bar{x}, \bar{v})$ est un point de selle du lagrangien signifie :

a) $\bar{x}$ maximise $l(., \bar{v})$

Par conséquent, parmi toutes les courbes $\phi(x,.)$ $(x \in X)$ la plus grande valeur prise en σ par une c-minorante affine de $\phi(x,.)$ définie par $\bar{v}$ est minimale pour $\bar{x}$.

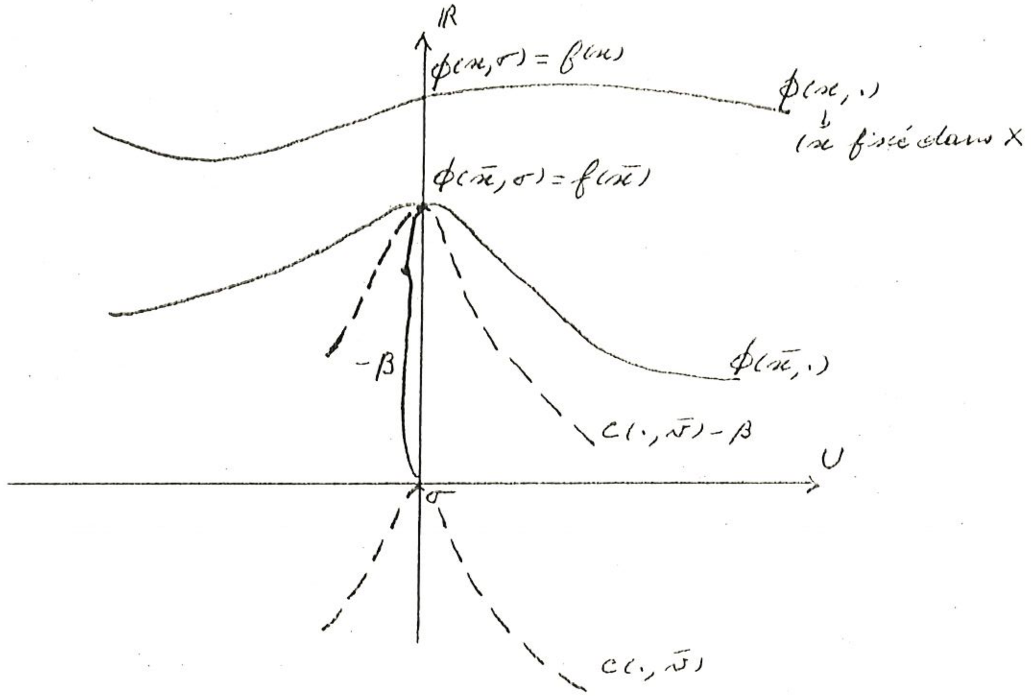
b) $\bar{v}$ minimise $l(\bar{x}, .)$

Il s'ensuit que la valeur de l'enveloppe supérieure des c-minorantes affines de $\phi(\bar{x},.)$ en σ est atteinte par une c-minorante affine de $\phi(\bar{x},.)$ définie par $\bar{v}$.

Par ailleurs, $-\beta$ correspond à la valeur prise en σ par l'enveloppe supérieure des c-minorantes de h

$(h(u) \overset{\text{déf}}{=} \inf\limits_{x \in X} \phi(x,u) \quad \forall u \in U)$.

En posant $U = \mathbb{R}$ nous pouvons plus ou moins nous représenter la situation par le dessin suivant.



Par conséquent, grâce aux hypothèses sur β, c et ϕ qui nous assurent que, en σ, $\phi(\bar{x},.)$ coïncidera avec l'enveloppe supérieure de ses c-minorantes affines (cfr interprétation du théorème II.1), $\bar{x}$ et $\bar{v}$ sont solutions respectivement de (P) et de (Q).

<u>Théorème II.7</u>

Si h est c-sous différentiable en σ alors pour toute solution $\bar{x}$ de (P) et toute solution $\bar{v}$ de (Q), le couple $(\bar{x},\bar{v})$ est un point de selle du lagrangien.
En particulier, pour $\bar{v}$ appartenant à c-$\partial h(\sigma)$ nous avons :

$$\sup_{x \in X} l(x,\bar{v}) = -\alpha$$

(démonstration voir annexe III)

Le théorème II.7 nous donne en quelque sorte une réciproque du théorème précédent.
Il nous dit que, si une e-minorante affine de h atteint la valeur finie $h(\sigma) = \alpha$ en σ alors tout couple de solutions primal-dual est un point de selle du lagrangien.

Chapitre III : Application de la théorie non convexe de la dualité aux problèmes de programmation mathématique.

Dans ce chapitre, nous étudions le cas particulier où le problème primal (P) est un programme mathématique. (voir définition I.12)

Comme dans le cas convexe, nous posons $Z = U$ et $\sigma = 0_Z$, et nous considérons la famille des problèmes perturbés $\{(P_u) \mid u \in U\}$ où,

$$(P_u) \qquad \inf_{x \in X} \{ f_0(x) \mid G(x) \leq u \} \qquad u \in U$$

Ceci revient à définir la fonctionnelle ϕ de $\bar{\mathbb{R}}^{X \times U}$ par

$$\phi(x,u) = \begin{cases} f_0(x) & \text{si } G(x) \leq u \\ +\infty & \text{sinon} \end{cases} \qquad x \in X,\ u \in U$$

et nous avons bien :

$$\phi(x,\sigma) = f(x) \overset{\text{déf}}{=} \begin{cases} f_0(x) & \text{si } G(x) \leq \sigma \\ +\infty & \text{sinon} \end{cases}$$

Suite à leur structure particulière, les problèmes de programmation mathématique ont des propriétés spécifiques que nous étudions dans les paragraphes qui suivent.

III.1. Stabilité du problème

Avant de donner un critère de stabilité pour (P), remarquons la propriété suivante :

Propriété III.1

Pour tout élément x de X, la fonctionnelle $\phi(x,.)$ de $\mathbb{R}^U$ est s.c.i en σ

Démonstration

Soit x un élément arbitraire de X et considérons un réel k strictement inférieur à $\phi(x,\sigma)$.
Nous allons démontrer qu'il existe un voisinage V de σ tel que :

$$k < \phi(x,u) \qquad \forall u \in V \qquad (1)$$

Nous envisageons deux cas :

1) $G(x) \leq \sigma$
Dans ce cas, $\phi(x,\sigma) = f_0(x)$ et n'importe quel voisinage de σ convient.
En effet, considérons un voisinage arbitraire V de σ.
Soit u un élément quelconque de V. Deux possibilités peuvent se présenter :

a) $G(x) \leq u$ ce qui entraîne : $\phi(x,u) = f_0(x) > k$

b) $G(x) \nleq u$ ce qui entraîne : $\phi(x,u) = +\infty > k$

Par conséquent, $k < \phi(x,u) \qquad \forall u \in V$.

2) $G(x) \not\leq \sigma$

Il s'ensuit : $\phi(x, \sigma) = +\infty$

Posons V égal au complémentaire de $\{u \in U \mid G(x) \leq u\}$

Nous allons prouver que V est un ouvert contenant σ, donc un voisinage de σ et que (1) est vérifiée par V.

Considérons $s : U \to U$ définie par $s(u) = G(x) - u$.

Puisque U est un espace vectoriel topologique, s est continue.

Or, $\{u \in U \mid G(x) \leq u\} = s^{\leftarrow}(\{u \in U \mid u \leq \sigma\}) = s^{\leftarrow}(C)$

Donc $\{u \in U \mid G(x) \leq u\}$ est un fermé car il est l'image réciproque du cône fermé C par une fonction continue.

Par conséquent, V est un ouvert contenant σ (cfr $G(x) \not\leq \sigma$)

En outre, nous obtenons suite à la définition de V

$$\phi(x, u) = +\infty \qquad \forall u \in V$$

L'inégalité (1) est donc vérifiée dans tous les cas et par conséquent, $\phi(x, .)$ est s.c.i en σ pour tout x dans X.

Nous allons énoncer maintenant un résultat de stabilité similaire au théorème I.5 pour un programme mathématique non spécialement convexe.

Théorème III.1

Dans le cas d'un programme mathématique, les trois conditions suivantes sont suffisantes pour assurer l'inf-stabilité du problème.

1) f_0 est s.c.i sur X

2) G est fermé pour l'ordre c'est à dire si nous considérons deux systèmes dirigés arbitraires, soient $\{u_\nu\}$ dans U et $\{x_\nu\}$ dans X, indexés de la même façon et convergeant respectivement vers u_0 appartenant à U et x_0 appartenant à X alors,

$G(x_\nu) \leq u_\nu$ pour tout élément ν de l'ensemble d'indices implique $G(x_0) \leq u_0$.

3) Il existe $\bar{u}$ appartenant à l'intérieur de C et un réel $\bar{\alpha}$ tels que :

- $\bar{\alpha} > \alpha$
- $\{x \in X \mid f_0(x) \leq \bar{\alpha}\} \cap \{x \in X \mid G(x) \leq -\bar{u}\}$ est compact
- $\{u \in U \mid \exists x \in X$ vérifiant $G(x) \leq u$ et $f_0(x) \leq \bar{\alpha}\}$ est un voisinage de σ.

(Si l'intérieur de C est vide, il suffit de demander l'existence d'un réel $\bar{\alpha}$ tel que :

- $\bar{\alpha} > \alpha$
- $\{x \in X \mid f_0(x) \leq \bar{\alpha}\}$ est compact
- $\{u \in U \mid \exists x \in X$ vérifiant $G(x) \leq u$ et $f_0(x) \leq \bar{\alpha}\}$ est un voisinage de σ

(démonstration voir annexe IV).

III.2 Etude du lagrangien d'un programme mathématique

Propriété III.2

Le lagrangien d'un programme mathématique a la forme suivante :

$$l(x,v) = -f_0(x) + \bar{c}(G(x), v) \qquad x \in X,\ v \in V \qquad (1)$$

où $\bar{c}$ est une fonctionnelle de $\bar{\mathbb{R}}^{U \times V}$ définie par :

$$\bar{c}(u,v) \overset{\text{déf}}{=} \operatorname{Sup} \{ c(u',v) \mid u' \in U,\ u \leq u' \} \qquad u \in U,\ v \in V$$

En outre, pour tout élément v de V, $\bar{c}(.,v)$ est décroissante.

Démonstration

Nous avons défini le lagrangien de la façon suivante :

$$l(x,v) = \operatorname*{Sup}_{u \in U} [c(u,v) - \phi(x,u)] \qquad x \in X,\ v \in V$$

En tenant compte de l'expression particulière de ϕ, nous obtenons :

$$l(x,v) = \max \{ \operatorname*{Sup}_{\substack{u \in U \\ G(x) \leq u}} [c(u,v) - f_0(x)],\ \operatorname*{Sup}_{\substack{u \in U \\ G(x) \nleq u}} [c(u,v) - \infty] \}$$

Puisque c est une fonctionnelle à valeurs réelles,

$$l(x,v) = \operatorname*{Sup}_{\substack{u \in U \\ G(x) \leq u}} [-f_0(x) + c(u,v)]$$

$$= -f_0(x) + \bar{c}(G(x), v)$$

Nous avons donc établi la formule (1).

Considérons un élément v arbitraire dans V, nous devons prouver que $\bar{c}(.,v)$ est décroissante sur U.

Soient u_1, u_2 appartenant à U tels que : $u_1 \leq u_2$

or, $\bar{c}(u, v) \overset{\text{déf}}{=} \operatorname{Sup} \{ c(u', v) \mid u' \in U, u \leq u' \}$ $u \in U, v \in V$

et $\{u' \in U \mid u_1 \leq u'\}$ contient $\{u' \in U \mid u_2 \leq u'\}$.

Par conséquent,

$$\bar{c}(u_1, v) \geq \bar{c}(u_2, v).$$

Propriété III.3

Soit c pointue en σ.

Si il existe un élément $\bar{v}$ de V tel que $\bar{c}(G(x), \bar{v})$ est finie alors, le lagrangien du programme mathématique (P) est symétrique c'est à dire :

$$-f(x) = \inf_{v \in V} l(x, v) \qquad x \in X \qquad (1)$$

$$g(v) = \sup_{x \in X} l(x, v) \qquad v \in V \qquad (2)$$

Démonstration

L'égalité (2) a déjà été vérifiée dans le cadre général ainsi que l'inégalité suivante :

$$-f(x) \leq \inf_{v \in V} l(x, v) \qquad \forall x \in X$$

Par conséquent, il suffit de prouver que :

$$-f(x) \geq \inf_{v \in V} l(x, v) \qquad \forall x \in X \qquad (3)$$

Suite à l'expression particulière du lagrangien nous avons :

$$\inf_{v \in V} l(x, v) = \inf_{v \in V} (-f_0(x) + \bar{c}(G(x), v)) = -f_0(x) + \inf_{v \in V} \bar{c}(G(x), v) \qquad (\forall x \in X)$$

Soit x, un élément arbitraire de X, deux cas peuvent se présenter :

1) $G(x) \leq \sigma$

Dans ce cas, nous obtenons :

$$\inf_{v \in V} \bar{C}(G(x), v) = \inf_{v \in V} \operatorname{Sup} \{ C(u', v) \mid u' \in U, G(x) \leq u' \} \leq 0$$

En effet, considérons un voisinage arbitraire N de σ.

Suite au caractère pointu de C en σ nous avons :

$\forall \varepsilon > 0$, $\exists v \in V$ et N' voisinage de σ, $N' \subset N$

tels que :

$$C(u, v) \leq \varepsilon \qquad \forall u \in N' \qquad (4)$$

$$C(u, v) \leq C(u, \bar{v}) + \eta \qquad \forall u \notin N' \qquad (5)$$

où nous supposons avoir préalablement défini

η par : $\quad \eta = -\bar{C}(G(x), \bar{v}) + \varepsilon$.

Soit u' un élément arbitraire de U vérifiant $G(x) \leq u'$, deux cas sont possibles :

a) $u' \in N'$ et donc, par (4), $\quad C(u', v) \leq \varepsilon$

b) $u' \notin N'$, il s'ensuit par (5)

$$C(u', v) \leq C(u', \bar{v}) - \bar{C}(G(x), \bar{v}) + \varepsilon$$

pour conséquent, vu la définition de $\bar{C}(G(x), \bar{v})$

$$C(u', v) \leq \varepsilon.$$

d'où, $\forall u' \in U$ tel que $G(x) \leq u'$ $\quad C(u', v) \leq \varepsilon$

ce qui implique : $\quad \bar{C}(G(x), v) \leq \varepsilon$

D'une façon générale nous obtenons :

$\forall \varepsilon > 0 \quad \exists v \in V \quad$ tel que $\quad \bar{C}(G(x), v) \leq \varepsilon$.

Nous en tirons la thèse pour les x satisfaisant $G(x) \leq \sigma$.

2) $G(x) \notin \sigma$.

Dans ce cas, nous devons démontrer que : $\inf\limits_{v \in V} \bar{c}(G(x), v) = -\infty$

Ce qui peut encore s'exprimer par :

$\forall M > 0 \quad \exists v \in V$ tel que $\forall u' \geq G(x) \quad c(u', v) \leq -M$

Soit M un réel positif arbitraire et considérons $N = \{u' / G(x) \leq u'\}^c$ et $\eta = -\bar{c}(G(x), \bar{v}) - M$

Nous avons déjà prouvé dans le cadre de la démonstration de la propriété III.1 que N est un voisinage de σ moyennant l'hypothèse $G(x) \notin \sigma$.

Par conséquent, suite au caractère pointu de c en σ, pour un réel ε strictement positif arbitraire il existe un voisinage N' de σ inclus dans N et un élément v de V tels que :

$$c(u, v) \leq \varepsilon \qquad \forall u \in N' \qquad (6)$$
$$c(u, v) \leq c(u, \bar{v}) + \eta \qquad \forall u \notin N' \qquad (7)$$

Par définition de N', N, η et par (7) nous avons donc :

$c(u, v) \leq c(u, \bar{v}) - \bar{c}(G(x), \bar{v}) - M \qquad \forall u \in U$ tel que $G(x) \leq u$

Il s'ensuit, par définition de $\bar{c}(G(x), \bar{v})$ que :

$\forall u \in U$ vérifiant $G(x) \leq u \qquad c(u, v) \leq -M$

Par conséquent, la thèse est bien vérifiée.

Dans tous les cas possibles pour les éléments x de X, l'inégalité (3) est donc satisfaite.

Reconsidérons la présentation générale d'un programme mathématique. Nous pouvons le redéfinir sous la forme d'un problème de minimisation contraint.
Soit, minimiser la fonction objective f_0 sous des contraintes définies par G ($G(x) \leq 0$).

Grâce à cette nouvelle présentation, nous retrouvons une certaine similitude entre notre lagrangien et des lagrangiens rencontrés usuellement.
En effet, notre lagrangien s'exprime comme la somme de la fonction objective et d'une fonction $\bar{c}$ des contraintes (représentées par G) et des multiplicateurs (représentés par v)

En fait, nous allons montrer que, en précisant les fonctionnelles c et G, nos lagrangiens se ramènent aux lagrangiens généralisés de la méthode de la fonction de pénalisation extérieure. (vue générale de la méthode voir annexe V.)

Nous allons reprendre les fonctionnelles couplantes données dans les exemples 1.a et 2.a et calculer la fonction $\bar{c}$ dans ces deux cas particuliers.

<u>Exemple 1.b</u>:

Soient $U = \mathbb{R}^m$, $V = \mathbb{R}^m_+$ et

$$c(u,v) = -\sum_{i=1}^{m} v_i \, |u_i| \quad , u = (u_1, u_2 \ldots u_m) \in \mathbb{R}^m$$
$$v = (v_1, v_2 \ldots v_m) \in \mathbb{R}^m_+$$

Nous avons démontré (cfr exemple 1.a) le caractère pointu de c en 0, l'origine de $\mathbb{R}^m$.

Considérons maintenant le programme mathématique défini par :

- $G = (g_1, \ldots g_p, g_{p+1}, \ldots g_m)$ où $1 \leq p \leq m$ et

$g_i : X \to \mathbb{R} \qquad i = 1, \ldots m$

- $C = \mathbb{R}^p_- \times \{0_{\mathbb{R}^{m-p}}\}$ où $0_{\mathbb{R}^{m-p}}$ désigne l'origine de $\mathbb{R}^{m-p}$

Dans ce cas, l'expression de $\bar{c}$ est la suivante :

$$\bar{c}(G(x), v) = -\sum_{i=1}^{p} \max(g_i(x), 0) - \sum_{i=p+1}^{m} v_i \, |g_i(x)| \qquad (x \in X, \; v \in V).$$

(démonstration voir annexe VI)

Pour interpréter la formulation particulière de $\bar{c}$ nous allons reformuler le programme mathématique dont nous sommes partis.

Soit, $\begin{cases} \inf\limits_{x \in X} f_0(x) \\ \\ \text{sous la contrainte : } G(x) \leq \sigma \quad (\sigma \text{ étant l'origine de } \mathbb{R}^m) \end{cases}$

Remarquons que, suite à la définition particulière du cône C, la contrainte résume en fait deux séries de contraintes.

1°) p contraintes d'inégalité :
$(g_1(x) - 0, \ldots, g_p(x) - 0) \in \mathbb{R}^m_-$ ce qui implique :
$g_1(x) \leq 0, \; g_2(x) \leq 0 \ldots, g_p(x) \leq 0$.

2°) $m-p$ contraintes d'égalité

$(g_{p+1}(x)-0, \ldots g_m(x)-0) = 0_{\mathbb{R}^{m-p}}$ c'est à dire :

$$g_{p+1}(x) = \cdots = g_m(x) = 0$$

Le lagrangien de ce programme mathématique s'écrit :

$$l(x,v) = -f_0(x) - \underbrace{\sum_{i=1}^{p} v_i \max(g_i(x),0)}_{(1)} - \underbrace{\sum_{i=p+1}^{m} v_i \, |g_i(x)|}_{(2)}$$

$(x \in X,\ v \in V)$

Nous y retrouvons la fonction objective à minimiser (f_0) et deux fonctions l'une dépendant des contraintes d'inégalité (cfr (1)) et l'autre des contraintes d'égalité (cfr (2)).

Les composantes v_i du vecteur v peuvent être assimilées à des constantes de pénalisation.
En effet, les fonctions multipliant chacune de celles-ci vérifient toutes les hypothèses pour être des fonctions de pénalisation. (voir annexe V).

Exemple 2.b

Soient $U = \mathbb{R}^m$, $V = \mathbb{R}_+ \times \mathbb{R}^m$ et

$c(u,v) = -v_0 |u|^2 - \langle w,u \rangle \quad , v = (v_0, w) \in \mathbb{R}_+ \times \mathbb{R}^m, u \in \mathbb{R}^m$

($|.|$ désigne la norme euclidienne dans $\mathbb{R}^m$ et $\langle .,. \rangle$ le produit scalaire usuel dans $\mathbb{R}^m$.)

Dans l'exemple 2.a, nous avions démontré le caractère pointu de c sur $\mathbb{R}^m$.

Reprenons maintenant le programme mathématique étudié dans l'exemple précédent.

Dans ce cas, l'expression de $\bar{e}$ est la suivante :

$$\bar{e}(G(x), v) = -\sum_{i=1}^{p}\left[v_0 \max^2\left(g_i(x), -\frac{w_i}{2v_0}\right) + w_i \max\left(g_i(x), -\frac{w_i}{2v_0}\right)\right]$$
$$-\sum_{i=p+1}^{m}\left[v_0 (g_i(x))^2 + w_i\, g_i(x)\right]$$
$$(x \in X,\ v = (v_0, w) \in R_+ \times \mathbb{R}^m)$$

(démonstration voir annexe VI)

Par conséquent,

$$l(x,v) = -f_0(x) \underbrace{-\sum_{i=1}^{p}\left[v_0 \max^2\left(g_i(x), -\frac{w_i}{2v_0}\right) + w_i \max\left(g_i(x), -\frac{w_i}{2v_0}\right)\right]}_{(1)}$$
$$\underbrace{-\sum_{i=p+1}^{m}\left[v_0 (g_i(x))^2 + w_i\, g_i(x)\right]}_{(2)}$$
$$(x \in X,\ v = (v_0, w) \in \mathbb{R}_+ \times \mathbb{R}^m)$$

Comme précédemment, nous retrouvons dans le lagrangien une partie relative aux contraintes d'inégalité (cfr (1)) et une autre relative aux contraintes d'égalité (cfr (2)).

Nous remarquons ici que nous sommes très proches du lagrangien augmenté de la méthode des multiplicateurs de Hestenes et Powell (résumé de la méthode voir annexe V.) le v_0 peut être assimilé à une constante de pénalisation et les w_i jouent le rôle des multiplicateurs au sens habituel.

Nous voyons à travers ces exemples que la méthode de Bälder est très générale.
En effet, suivant le choix de la fonctionnelle couplante c nous pouvons obtenir des lagrangiens correspondant à différentes méthodes connues où certains de nos multiplicateurs s'interprètent comme des constantes de pénalisation et les autres comme des multiplicateurs au sens habituel.

Après avoir montré la correspondance entre les lagrangiens de cette théorie et les lagrangiens généralisés de la méthode de pénalisation extérieure, nous allons exposer dans le paragraphe suivant quelques théorèmes de convergence.

III.3. Résultats de convergence pour la méthode de pénalisation extérieure

Nous allons à l'avenir considérer les éléments de V comme des multiplicateurs.
Nous désignerons par le terme de MULTIPLICATEUR EXACT tout élément $\bar{v}$ de V qui permet en une seule maximisation du lagrangien d'atteindre la valeur optimale du problème initial (P)
c'est à dire :

$$\bar{v} \text{ multiplicateur exact ssi } \sup_{x \in X} l(x, \bar{v}) = -\alpha$$

Le théorème II.7 nous avait déjà donné une condition suffisante pour qu'un élément de V soit un multiplicateur exact.

En général, nous nous intéressons beaucoup plus aux solutions optimales (c'est à dire aux points $\bar{x}$ de X tels que $f(\bar{x}) = \alpha$) qu'à la valeur optimale α.
Par conséquent, nous voudrions, pour un multiplicateur exact, que la solution optimale de la maximisation du lagrangien soit solution optimale du problème (P) ; ce qui n'est pas spécialement immédiat.
Nous verrons, ci-après, un théorème donnant des hypothèses suffisantes pour que cela se produise.

Pour atteindre les solutions optimales de (P), nous pouvons procéder également par une maximisation séquentielle du lagrangien.

Pour cela, nous engendrons une suite de multiplicateurs $\{v_k\}_{k\in\mathbb{N}}$ telle que :

$$\lim_{k\to\infty} \sup_{x\in X} l(x, v_k) = \lim_{k\to\infty} g(v_k) = -\alpha$$

Les v_k sont dans ce cas "presque" c-sous-différentiels de h. En effet, vu l'interprétation géométrique de $h^c = g$,

$g(v_k)$ tend vers $(-\alpha)$ signifie que, pour k croissant, la valeur maximale prise en σ par une c-minorante affine de h définie par v_k (soit $-g(v_k)$) tend vers α.

Nous pourrions donc poser $g(v_k) = -\alpha + \varepsilon_k$ (1)

où $\{\varepsilon_k\}_{k\in\mathbb{N}}$ est une suite de nombres réels positifs tendant vers 0.

Par définition de g, nous avons :

$h(u) - h(\sigma) \geq c(u, v_k) - g(v_k) - h(\sigma) \qquad \forall u\in U, \forall k\in\mathbb{N}$

Par conséquent, vu (1),

$h(u) - h(\sigma) \geq c(u, v_k) - \varepsilon_k \qquad \forall u\in U, \forall k\in\mathbb{N}$

Comme c est normalisée en σ, ceci est équivalent à :

$v_k \in c\text{-}\partial_{\varepsilon_k} h(\sigma) \qquad \forall k\in\mathbb{N}$

Pour cette méthode, nous espérons que la suite des solutions optimales x_k des maximisations successives du lagrangien en v_k convergera vers une solution optimale du problème primal.

Cependant, tout comme pour la méthode des multiplicateurs exacts cela n'est pas immédiat.

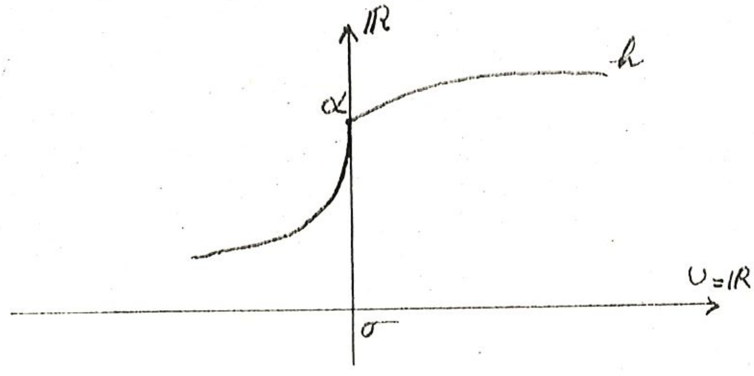
Par exemple, il est parfois nécessaire que les v_k engendrés ou du moins une de leurs composantes si ils appartiennent à $\mathbb{R}^n$ tendent vers l'infini pour k croissant. Ceci provoque invariablement un mauvais conditionnement des matrices hessiennes des lagrangiens généralisés et par conséquent, la recherche d'un maximisant x_k de $l(\cdot, v_k)$ devient de plus en plus pénible pour k croissant.

En effet, si la matrice hessienne associée au problème non contraint $\sup_{x \in X} l(x, v_k)$ a des valeurs propres trop différentes, les méthodes habituelles (méthode du gradient, ..) ont des vitesses de convergence très faibles (cfr [5])

Nous avions déjà abordé de tels problèmes dans l'annexe V dans le cadre de la méthode de pénalisation pure.

Pour cette théorie plus élaborée, l'obligation de faire croître certaines composantes des multiplicateurs peut paraître moins naturelle que pour la méthode de pénalisation pure. Cependant, nous pouvons concevoir certains cas où c'est nécessaire.

Par exemple, choisissons la fonctionnelle couplante de l'exemple 1.a en prenant $m = 1$ et supposons que la fonction de perturbation est tangente à gauche à l'axe des y en σ. Nous voyons immédiatement que v_k devra obligatoirement tendre vers $(+\infty)$ pour que $g(v_k)$ converge vers $(-\alpha)$. (cfr interprétation géométrique de $g = h^c$).



Un autre problème est la convergence de la suite $\{x_k\}_{k \in \mathbb{N}}$ vers une solution optimale de (P).
En effet, les points d'adhérence de cette suite ne sont pas nécessairement des solutions optimales de (P).
Nous verrons également un théorème de convergence pour cette méthode.

Mais avant d'énoncer ces théorèmes de convergence nous avons besoin d'une propriété complémentaire pour la fonctionnelle couplante.

<u>Définition III.1</u>

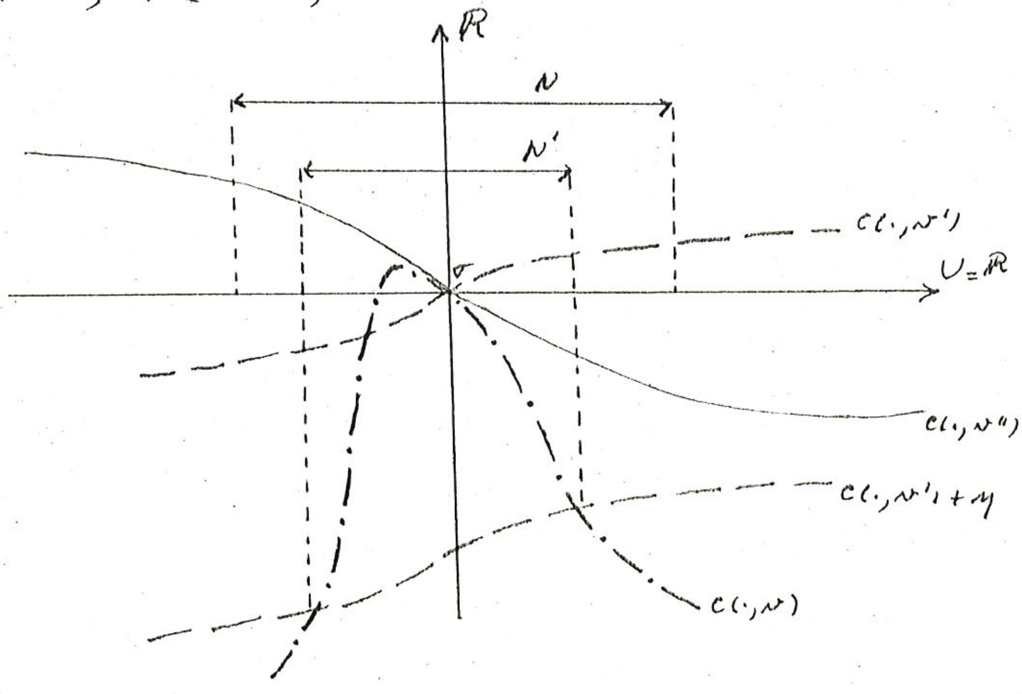
Une fonctionnelle couplante c de $\mathbb{R}^{U \times V}$ est dite FLEXIBLE en σ si, $\forall N$ voisinage de σ

$\forall v', v'' \in V$

$\forall \eta \in \mathbb{R}$

$\exists v \in V$ et $\exists N'$ voisinage de σ, $N' \subseteq N$ tels que :

$$\begin{cases} c(u,v) \leq c(u,v') + \eta & \forall u \notin N' \quad (III.1) \\ c(u,v) \leq c(u,v'') & \forall u \in N' \quad (III.2) \end{cases}$$



Remarquons que, de nouveau, nous retrouvons une interprétation géométrique très similaire à celle des fonctions pointues en σ. Cependant, ce n'est pas exactement la même chose. En fait, nous n'avons pas d'équivalence, ni même d'implication entre les deux notions bien qu'elles paraissent assez fortement liées.
La proposition suivante permet cependant de faire un joint entre celles-ci.

<u>Proposition III.1</u>

Si il existe v_0 appartenant à V tel que $c(\cdot, v_0)$ est s.c.s en σ alors la flexibilité de c en σ implique le caractère pointu en σ.

<u>Démonstration</u>

Appliquons la définition de flexibilité en fixant $v'' = v_0$
Par conséquent, pour tout v' dans V, pour tout voisinage N de σ et tout réel η, il existe un élément v de V et un voisinage N'' de σ inclus dans N tels que :

$$c(u, v) \leq c(u, v_0) \qquad \forall u \in N''$$
$$c(u, v) \leq c(u, v') + \eta \qquad \forall u \notin N''$$

Comme $c(\cdot, v_0)$ est s.c.s en σ, pour tout réel ε strictement positif, il existe un voisinage $\bar{N}_\varepsilon$ de σ tel que

$$c(u, v_0) \leq \varepsilon \qquad \forall u \in \bar{N}_\varepsilon$$

Posons $N' = N'' \cap \bar{N}_\varepsilon$
Par conséquent, $\forall \varepsilon > 0$, $\forall v' \in V$, $\forall N$ voisinage de σ et $\forall \eta \in \mathbb{R}$

$\exists v \in V$, $\exists N'$ voisinage de σ, $N' \subset N$ tels que :

$$\left| \begin{array}{ll} c(u, v) \leq \varepsilon & \forall u \in N' \\ c(u, v) \leq c(u, v') + \eta & \forall u \notin N' \end{array} \right.$$

<u>Remarque</u> :

les fonctionnelles couplantes des exemples 1.a et 2.a sont toutes deux flexibles en σ.
(démonstration voir annexe VII).

Supposons que nous avons construit une suite de paramètres duaux $\{v_k\}_{k \in \mathbb{N}}$ ajustée de telle sorte que la suite $\{g(v_k)\}_{k \in \mathbb{N}}$ converge vers β qui est supposé fini et égal à $-\alpha$.

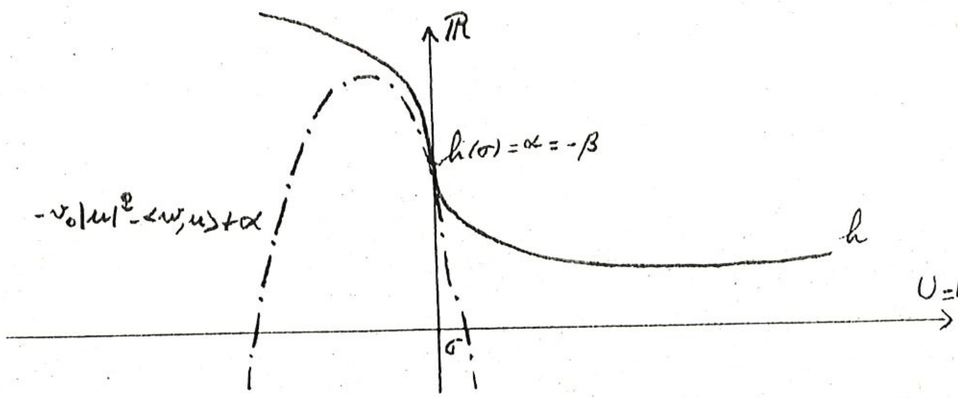
Remarquons qu'intuitivement un tel ajustement se conçoit très bien.

Par exemple,

a) Si nous considérons la fonctionnelle couplante de l'exemple 1.a, en faisant tendre vers l'infini les composantes de v, $g(v)$ convergera alors vers β.
En effet, pour v croissant la fonctionnelle devient de plus en plus pointue en σ ce qui entraîne (moyennant quelques hypothèses sur h cfr théorème II.1) la convergence de $g(v_k)$ vers β. (cfr interprétations géométriques de g et de β.)

b) Si nous envisageons la fonctionnelle couplante de l'exemple 2.a, nous allons jouer sur les deux paramètres (v_0 et w) d'une façon similaire à l'ajustement de la constante de pénalisation et des multiplicateurs dans la méthodes des multiplicateurs de Hestenes et Powell.
(cfr annexe V)

Remarquons que si nous fixons $w=0$, nous retrouvons le même ajustement que pour l'exemple 1.
Pour des fonctions de perturbation d'un certain type (par exemple, admettant presque un point d'inflexion à tangente verticale en σ dans le cas où $U=\mathbb{R}$) il est intéressant de pouvoir jouer sur deux paramètres. Cela permet d'éviter, dans certains cas, de faire tendre un des paramètres vers l'infini et donc d'éviter les ennuis numériques qui s'ensuivent. (cfr annexe V.)



Pour la fonction de perturbation représentée sur le dessin, si on emploie la fonctionnelle couplante de l'exemple 1, il faudrait que v soit très grand (cela correspond à une pointe très fine) pour que :

$$c(u,v)+\alpha \leq h(u) \qquad \forall u \in U \qquad (1)$$

Par contre, si on considère la fonctionnelle couplante de l'exemple 2, l'inégalité (1) est valable pour v_0 et w nettement plus petit.
Remarquons que pour cette fonctionnelle couplante, "fixer v_0" correspond à choisir une certaine largeur de l'aiguille parabolique et faire varier w correspond à translater celle-ci.
Donc, chaque étape de l'ajustement consistera d'abord

à se fixer v_0, ensuite à choisir un élément w rendant $g(v)$ le plus proche possible de $(-\alpha)$ (cfr $v=(v_0, w)$) avant d'augmenter v_0 (ce qui correspond à affiner l'aiguille parabolique).

Revenons à l'aspect théorique du problème.
Nous supposons donc avoir construit une suite $\{v_k\}_{k\in N}$ telle que $\{g(v_k)\}_{k\in N}$ converge vers $\beta=-\alpha$ c'est à dire une suite de presque c-sous-différentiels de h (cfr page 90)

$\forall k \in N \quad v_k \in c\text{-}\partial_{\varepsilon_k} h(\sigma)$ où $\{\varepsilon_k\}_{k\in N}$ est une suite de réels positifs convergeant vers zéro.

Faisons l'hypothèse supplémentaire que chaque pas d'optimisation de $l(.,v_k)$ donne un "presque" optimisant x_k avec une certaine précision λ_k c'est à dire :

$$l(x_k, v_k) \geq \sup_{x\in X} l(x, v_k) - \lambda_k \quad , \forall k \in N$$

Sous ces hypothèses, nous pouvons maintenant énoncer un théorème donnant des conditions suffisantes pour que tout point d'adhérence de la suite $\{x_k\}_{k\in N}$ soit une solution optimale du problème initial (P).

Théorème III.2

Supposons que la topologie sur V est engendrée par une norme et que la fonctionnelle couplante c de $R^{U\times V}$ est flexible en σ.

Soit $\{N_k\}_{k\in\mathbb{N}}$ une suite de boules ouvertes autour de σ dont les diamètres décroissent vers zéro.

Afin de résoudre le problème initial, nous supposons avoir engendré une suite $\{v'_k\}_{k\in\mathbb{N}}$ dans V telle que

$$v'_k \in c\text{-}\partial_{\varepsilon_k} h(\sigma) \quad , k \in \mathbb{N}$$

où $\{\varepsilon_k\}_{k\in\mathbb{N}}$ est une suite de réels positifs convergeant vers zéro.

Soient $\{\lambda_k\}_{k\in\mathbb{N}}$ et $\{\eta_k\}_{k\in\mathbb{N}}$ deux suites de réels telles que :

$$\begin{cases} \lambda_k \geq 0 & \forall k \in \mathbb{N} \\ \eta_k < -\varepsilon_k - \lambda_k & \forall k \in \mathbb{N} \end{cases}$$

Pour $k \in \mathbb{N}$, considérons un élément v_k de V vérifiant les formules de flexibilité (III.1) et (III.2) pour $N = N_k$, $\eta = \eta_k$, $v' = v'' = v_k$.

Soit x_k appartenant à X tel que:

$$l(x_k, v_k) \geq \sup_{x\in X} l(x, v_k) - \lambda_k \qquad k \in \mathbb{N}$$

Alors, si G est fermé pour l'ordre, tout point d'adhérence de $\{x_k\}_{k\in\mathbb{N}}$ satisfait les contraintes.

Si, en outre, f_0 est s.c.i, $\{\eta_k\}_{k\in\mathbb{N}}$ et $\{\lambda_k\}_{k\in\mathbb{N}}$ convergent vers zéro et si la famille $\{c(\cdot, v_k)\}_{k\in\mathbb{N}}$ est semi-continue supérieurement en σ uniformément par rapport à k,

Alors, tout point d'adhérence de $\{x_k\}_{k\in\mathbb{N}}$ est une solution de (P)

(démonstration voir annexe VIII.)

Remarques :

1) La semi-continuité supérieure en σ uniforme par rapport à k de la famille $\{c(\cdot, \nu_k)\}_{k \in \mathbb{N}}$ est immédiatement vérifiée si : $c(\cdot, \nu_k) \leq 0 \qquad \forall k \in \mathbb{N}$.

En effet, il suffit de prouver que :

$\forall \varepsilon > 0$, $\exists V_\varepsilon$ voisinage de σ tel que :

$c(u, \nu_k) \leq c(\sigma, \nu_k) + \varepsilon \qquad \forall k \in \mathbb{N}, \ \forall u \in V_\varepsilon$.

Ceci est évident suite à l'hypothèse et à la normalisation de c en σ.

2) Nous pouvons également appliquer le critère suivant si $c(\cdot, \nu_k)$ est concave pour tout k dans $\mathbb{N}$.

Proposition III.2

Soit $\{a_i\}_{i \in I}$ une famille de fonctionnelles concaves sur l'espace vectoriel U équipé d'une norme notée $\|.\|$ à valeurs réelles.

Soit $a_i'(\sigma; u) = \lim\limits_{\substack{\lambda \to 0 \\ >}} \dfrac{a_i(\sigma + \lambda u) - a_i(\sigma)}{\lambda}$ la dérivée directionnelle de a_i en σ dans la direction u. ($u \in U$, $i \in I$)

Alors, $\{a_i\}_{i \in I}$ est semi-continue supérieurement en σ uniformément par rapport à i si, il existe une constante K telle que pour tout i dans I, pour tout u (appartenant à U) de norme unité nous avons :

$$a_i'(\sigma; u) \leq K.$$

(démonstration voir annexe IX)

Les fonctionnelles couplantes des exemples 1.a et 2.a sont telles que les familles $\{c(\cdot, v)\}_{v \in V}$ sont semi-continues supérieurement en o uniformément par rapport à v. (vérification voir annexe IX.)

Donnons quelques commentaires sur la signification du théorème III.2.

Ce théorème ne spécifie pas comment ajuster les paramètres duaux $\{v'_k\}$, il n'explique pas comment trouver, à partir de ceux-ci, les paramètres $\{v_k\}$ dont l'existence est assurée par l'hypothèse de flexibilité.

Son but est plutot de montrer comment utiliser une procédure de résolution du dual pour résoudre le primal dans le sens de la construction d'une suite $\{x_k\}_{k \in \mathbb{N}}$ qui, de façon asymptotique, satisfait les contraintes et minimise le primal.

A l'étude des remarques sur l'ajustement des paramètres duaux (page 94, 95) nous sentons que le choix des paramètres λ_k, η_k et N_k sera un compromis entre l'exactitude des calculs dans les étapes de maximisation, la quantité de croissance de quelques composantes des multiplicateurs et la vitesse à laquelle les points x_k approchent la région déterminée par les contraintes (cfr démonstration du théorème où l'on a ; $\forall k \in \mathbb{N} \quad G(x_k) \leq u_k, \ u_k \in N_k$)

Ce théorème de convergence est une généralisation d'un résultat établi par Rockafellar pour le cas particulier du Lagrangien de l'exemple 2.6 avec uniquement des contraintes d'inégalité. (cfr [7] p 274, théorème 3)

Rockafellar envisage le programme mathématique défini par :

. $U = \mathbb{R}^m$, $V = \mathbb{R}_+ \times \mathbb{R}^m$

. $c(u,v) = -v_0 |u|^2 - \langle w, u \rangle$; $v = (v_0, w) \in \mathbb{R}_+ \times \mathbb{R}^m$, $u \in \mathbb{R}^m$

. $G = (g_1, g_2, \dots g_m)$ où $g_i : X \to \mathbb{R}$ $i = 1, \dots m$

. $f_0 : X \to \mathbb{R}$ la fonction objective à minimiser

. $C = \mathbb{R}^m_-$, le cône négatif dans U

soit, (P) $\min\limits_{x \in X} f_0(x)$

sous contrainte : $G(x) \leq 0$ (0 étant l'origine de $\mathbb{R}^m$)

Nous gardons l'hypothèse $-\alpha = \beta < +\infty$.

Enoncé du théorème de Rockafellar

Considérons une suite $\{v_k\}_{k=1}^{\infty}$ telle que :

$\forall k \in \mathbb{N}$ $v_k = (v_{0_k}, w_k)$ $v_{0_k} \in \mathbb{R}_+$, $w_k \in \mathbb{R}^m$

et supposons l'existence d'un réel δ strictement positif tel que :

$$\lim_{k \to \infty} g(v'_k) = \beta < +\infty \qquad (1)$$

où $v'_k = (v_{0_k} - \delta, w_k)$ $\forall k \in \mathbb{N}$

Soit $\{x_k\}_{k=1}^{\infty}$ une suite d'éléments de X satisfaisant :

$$\forall k \in \mathbb{N}_0 \quad l(x_k, v_k) \geq \sup_{x \in X} l(x, v_k) - \alpha_k \qquad (2)$$

où $\{\alpha_k\}_{k=1}^{\infty}$ est une suite de nombres réels convergeant vers zéro.

Alors, la suite $\{x_k\}_{k=1}^{\infty}$ satisfait asymptotiquement les contraintes et,

$$\liminf_{k \to \infty} \frac{w_{ik}}{v_{0k}} \geq 0 \qquad i = 1..m$$

Si, en outre, $\{w_k\}_{k=1}^{\infty}$ est une suite bornée alors $\{x_k\}_{k=1}^{\infty}$ minimise asymptotiquement (P)

Définitions données par Rockafellar

. Une suite $\{x_k\}_{k=1}^{\infty}$ d'éléments de X satisfait asymptotiquement les contraintes si et seulement si :

$$\limsup_{k \to \infty} g_i(x_k) \leq 0 \qquad i = 1..m$$

. Une suite $\{x_k\}_{k=1}^{\infty}$ d'éléments de X minimise asymptotiquement (P) si elle satisfait asymptotiquement les contraintes et donne la valeur minimale possible (dans le domaine contraint) à $\limsup_{k \to \infty} f_0(x_k)$.

Remarques.

1) Comme nous avons adapté l'énoncé du théorème de Rockafellar au contexte dans lequel nous travaillons nous en donnons la démonstration dans l'annexe VIII.

2) D'après les définitions données par Rockafellar nous voyons que ce théorème donne un résultat légèrement plus faible que le théorème III.2.

Cependant, si nous ajoutons les hypothèses suivantes :

a) G fermé pour l'ordre

alors, en regardant la démonstration de Rockafellar nous voyons que tout point d'adhérence de $\{x_k\}_{k=1}^{\infty}$ satisfait les contraintes.

(b) f_0 s.c.i

nous aurons, en outre, que tout point d'adhérence de $\{x_k\}_{k=1}^{\infty}$ est une solution optimale de (P).

Nous allons vérifier que, mis à part ces deux hypothèses supplémentaires, toutes les hypothèses données par Bälder sont reprises sous une forme différente dans l'énoncé du théorème de Rockafellar.

. V est bien un espace normé et c une fonctionnelle flexible en o.

. La suite $\{v'_k\}_{k=1}^{\infty}$ vérifie $\lim_{k\to\infty} g(v'_k) = -\alpha$ et, par conséquent (cfr assertion page 90) il existe une suite $\{\varepsilon_k\}_{k=1}^{\infty}$ de nombres réels positifs convergeant vers zéro telle que : $v'_k \in c - \partial_{\varepsilon_k} h(o) \qquad \forall k \in \mathbb{N}_0$.

. La suite $\{\alpha_k\}_{k=1}^{\infty}$ joue le rôle de la suite $\{\lambda_k\}_{k=1}^{\infty}$

. Considérons une suite de nombres négatifs $\{\eta_k\}_{k=1}^{\infty}$ où $\eta_k < -\varepsilon_k - \alpha_k$ et une suite de boules ouvertes N_k de centre o et de rayon égal à $\sqrt{\frac{-\eta_k}{\delta}}$

Nous allons vérifier que v_k satisfait les formules ((III.1) et (III.2)) de flexibilité pour $v' = v'' = v'_k$, $N = N_k$ et $\eta = \eta_k$

En effet, posons $N' = N_k$.

Suite à la définition de c et de N_k, nous avons immédiatement :

. $c(u, v_k) = -v_{o_k} |u|^2 - \langle w_k, u \rangle \leq -(v_{o_k} - \delta)|u|^2 - \langle w_k, u \rangle = c(u, v'_k)$
$\forall u \in N'$

. $c(u, v_k) \leq c(u, v'_k) + \eta_k = -(v_{o_k} - \delta)|u|^2 - \langle w_k, u \rangle + \eta_k$
$\forall u \notin N'$

car u n'appartient pas à N' est équivalent à $|u|^2 > \frac{-\eta_k}{\delta}$

. Comme $\{\alpha_k\}$ et $\{\varepsilon_k\}$ convergent vers zéro nous pouvons choisir pour la suite $\{\eta_k\}_{k=1}^{\infty}$ une suite convergeant vers zéro et donc les diamètres des N_k décroissent vers zéro.

. En outre, x_k est un presque maximisant du lagrangien pour tout k.

. Le caractère borné de la suite $\{w_k\}_{k=1}^{\infty}$, implique la semi-continuité supérieure en o uniformément par rapport à k de la famille $\{c(\cdot, v_k)\}_{k=1}^{\infty}$. (cfr annexe IX)

Par conséquent, en analysant les hypothèses de Rockafellar et de Bälder une à une nous voyons la profonde similitude existant entre ces deux résultats.

Rockafellar donne également un exemple prouvant que l'hypothèse "$\{w_k\}_{k=1}^{\infty}$ suite bornée" est une condition indispensable pour obtenir la théorie et par conséquent, dans le cas plus général de Bäloder, l'hypothèse de semi-continuité supérieure uniforme par rapport à k de la famille $\{c_i, v_k\}_{k=1}^{\infty}$.

Voici l'exemple :

Considérons $S = \{ y \in \mathbb{R}^2 \mid y_1 \leq 0,\ y_1^2 + 2 y_2 \leq 0 \}$ et posons $X = \{ x \in \mathbb{R}^3 \mid x_1 y_1 + x_2 y_2 - x_3 \leq 0 \quad \forall (y_1, y_2) \in S \}$.
Nous pouvons vérifier de façon immédiate que X est un cône convexe fermé.

Soient $f_0 : \mathbb{R}^3 \to \mathbb{R} : x \to x_3$

$G = (g_1, g_2)$ telle que $\left| \begin{array}{l} g_1 : \mathbb{R}^3 \to \mathbb{R} : x \mapsto x_1 \\ g_2 : \mathbb{R}^3 \to \mathbb{R} : x \mapsto x_2 \end{array} \right.$

Le problème à résoudre est le suivant :

$$(P) \left\{ \begin{array}{l} \min\limits_{x \in X} f_0(x) \\ \\ \text{sous contraintes : } G(x) \leq \sigma \end{array} \right.$$

où σ est l'origine de $\mathbb{R}^3$ et $\leq$ est l'ordre défini par le cône convexe négatif dans $\mathbb{R}^3$, soit $\mathbb{R}^3_-$.

Remarquons que X peut encore s'écrire sous la forme :

$$X = \{ x \in \mathbb{R}^3 \mid x_3 \geq \Psi((x_1, x_2)) \}$$

où $\Psi : \mathbb{R}^2 \to \mathbb{R} : x \mapsto \sup \{ x_1 y_1 + x_2 y_2 \mid (y_1, y_2) \in S \}$

En fait, $\Psi((x_1, x_2)) = \begin{cases} \dfrac{x_1^2}{2x_2} & \text{si } x_1 \leq 0 \text{ et } x_2 > 0 \quad (1) \\ 0 & \text{si } x_1 \geq 0 \text{ et } x_2 \geq 0 \quad (2) \\ +\infty & \text{dans les autres cas.} \quad (3) \end{cases}$

(2) et (3) se justifient aisément en considérant les définitions de Ψ et de S, et en analysant les signes de x_1, x_2 et y_1, y_2.

Pour le cas (1), considérons la fonction à maximiser soit $\underbrace{x_1 y_1}_{a)} + \underbrace{x_2 y_2}_{b)}$ avec $\begin{cases} x_1 \leq 0, \ x_2 > 0 \\ y_1 \leq 0 \text{ et } y_1^2 + 2y_2 \leq 0 \end{cases}$

Le terme a) est toujours positif. Par contre, le terme b) est toujours négatif nous avons donc intérêt à choisir le plus petit y_2 possible soit : $y_2 = \dfrac{-y_1^2}{2}$

Nous obtenons alors la fonction "$x_1 y_1 - x_2 \dfrac{y_1^2}{2}$" (4)
avec $x_1 \leq 0$, $x_2 > 0$, $y_1 \leq 0$.

Calculons sa dérivée par rapport à y_1, nous obtenons :

$$x_1 - x_2 y_1$$

Cette fonction admettra donc un maximum en $y_1 = \dfrac{x_1}{x_2} \leq 0$

Par conséquent, en remplaçant dans (4) nous avons :

$$\Psi((x_1, x_2)) = \frac{x_1^2}{2x_2} \quad \text{pour } x_1 \leq 0, \ x_2 > 0.$$

Remarquons que Ψ est décroissante en x_1 pour x_2 fixé et en x_2 pour x_1 fixé.

Calculons la fonction de perturbation h et sa conjuguée g.

1) $h(u) \overset{\text{déf}}{=} \inf_{x \in X} \{ f_0(x) \mid G(x) \leq u \}$ $\quad u = (u_1, u_2) \in \mathbb{R}^2$

$$= \inf \{ x_3 \mid x_1 \leq u_1,\ x_2 \leq u_2,\ \Psi(x_1, x_2) \leq x_3 \}$$

Comme Ψ est décroissante en x_1 et en x_2 l'infimum est atteint pour $x_1 = u_1$, $x_2 = u_2$.

Par conséquent,

$$h(u) = \Psi((u_1, u_2)) \qquad \forall\, u = (u_1, u_2) \in \mathbb{R}^2.$$

2) $g(v) \overset{\text{déf}}{=} \sup_{u \in \mathbb{R}^2} \{ -v_0 |u|^2 - \langle w, u \rangle - h(u) \}$ $\quad v = (v_0, w) \in \mathbb{R}_+ \times \mathbb{R}^2$

Comme $h(0) = 0$ et $h\left(\frac{u}{v_0}\right) = \frac{1}{v_0} h(u)$ $\quad \forall u \in U,\ \forall v_0 > 0$
(cfr définition de Ψ)

nous avons :

$$\begin{cases} g(v) \leq 0 & \forall\, v \in V \qquad (5) \\ g((v_0, w)) = \dfrac{1}{v_0}\, g((1, w)) & \forall\, v_0 > 0,\ \forall w \in \mathbb{R}^2 \qquad (6) \end{cases}$$

Construisons une suite $\{v_k\}_{k=1}^{\infty}$ répondant aux hypothèses du théorème excepté le caractère borné pour $\{w_k\}_{k=1}^{\infty}$.

Pour tout entier naturel non nul k, posons :

$$u_k = (-k^{-1}, k^{-3}) \quad ; \quad x_k = \left(-k^{-1}, k^{-3}, \frac{k}{2}\right)$$
$$v_{0k} = 1 \quad ; \quad w_k = -\nabla h(u_k) - 2 u_k$$
$$v_k = (v_{0k}, w_k)$$

Calculons $g(v_k)$

$$g(v_k) = \sup_{u \in \mathbb{R}^2} \{ -|u|^2 - \langle -\nabla h(u_k) - 2u_k, u \rangle - h(u) \}$$

Suite au choix particulier de v_k ce supremum est atteint en u_k et par conséquent, $g(v_k) = -|u_k|^2$.

Il s'ensuit, par définition de u_k, que $\{g(v_k)\}_{k=1}^{\infty}$ converge vers zéro et donc, par (6) $\{g((v_{0_k} - \delta, w_k))\}_{k=1}^{\infty}$ converge également vers zéro pour tout δ appartenant à $]0,1[$

Nous vérifions donc bien l'hypothèse de convergence vers la valeur optimale $\beta = \inf_{v \in V} g(v)$ du problème dual. (cfr (5))

En outre, nous savons que $h(u_k) = \varphi(u_k) = \frac{k}{2}$ et par définition, $h(u_k) = \inf_{x \in X} \{ f_0(x) \mid G(x) \leq u_k \}$

l'infimum est donc atteint pour $x = x_k$.

Nous avons donc que x_k maximise $L(x, v_k)$.

En effet, le supremum dans $g(v_k)$ est atteint en u_k et l'infimum dans $h(u_k)$ est atteint en x_k.

Or, $g(v_k) = \sup_{x \in X} L(x, v_k)$

et $g(v_k) = \sup_{u \in \mathbb{R}^2} \{ -v_{0_k} |u|^2 - \langle w, u \rangle - h(u) \}$

d'où x_k maximise $L(x, v_k)$.

Toutes les hypothèses sont donc satisfaites, excepté le caractère borné de $\{w_k\}_{k=1}^{\infty}$ et nous obtenons:

$$f_0(x_k) = \frac{k}{2} \qquad \forall k \in \mathbb{N}_0.$$

$\{x_k\}_{k=1}^{\infty}$ ne minimise pas asymptotiquement (P) car la limite de $\{f_0(x_k)\}_{k=1}^{\infty}$ est $+\infty$.

. Considérons maintenant le cas d'un multiplicateur exact avec une optimisation exacte du lagrangien.

Supposons que d'une manière ou d'une autre (nous ne discutons pas ce problème maintenant), nous avons été capable de déterminer un multiplicateur exact $\bar{v} \in C_-\partial h(o)$ ce qui suppose implicitement que $-\alpha = \beta < \infty$.

Faisons l'hypothèse qu'il existe un élément $\bar{x}$ de X minimisant exactement le lagrangien correspondant; soit $l(\cdot, \bar{v})$.

La question qui se pose est de savoir si $\bar{x}$ est une solution optimale du problème original.

Le théorème suivant donne des conditions suffisantes pour que cela soit vrai.

Théorème III.3

Soit c appartenant à $\mathbb{R}^{U \times V}$ flexible en σ.
Supposons qu'il existe $\bar{v} \in c - \partial h(\sigma)$ et que le point $\bar{\bar{v}}$ correspondant à (III.1) et (III.2) (les formules de flexibilité pour un voisinage N de σ, un réel η strictement négatif et $v' = v'' = \bar{v}$) a la propriété additionnelle suivante :

$$c(u, \bar{\bar{v}}) < c(u, \bar{v}) \qquad \forall u \in N \setminus \{\sigma\}.$$

Soit $\tilde{x}$, un élément de X qui maximise $\ell(., \bar{\bar{v}})$ et supposons que $\operatorname{Sup} \{c(u, \bar{\bar{v}}) \mid u \in U,\ G(\tilde{x}) \leq u\}$ est atteint.
Alors $\tilde{x}$ est une solution optimale du problème initial (P)

(démonstration voir annexe X)

Remarques :

1) Remarquons que l'hypothèse sur $\operatorname{Sup} \{c(u, \bar{\bar{v}}) \mid u \in U, G(\tilde{x}) \leq u\}$ est réalisée dans le cas particulier où :

- U est un espace vectoriel topologique de dimension finie
- $c(., \bar{\bar{v}})$ est concave
- $\{u \mid c(u, \bar{\bar{v}}) \geq 0\}$ est un compact non vide.

Démonstration

En premier lieu, constatons que sous ces hypothèses les ensembles niveaux de $c(., \bar{\bar{v}})$ sont compacts.
(Par ensemble niveau de $c(., \bar{\bar{v}})$ nous désignons tout ensemble du type $\{u \in U \mid c(u, \bar{\bar{v}}) \geq \lambda\}$, $\lambda \in \mathbb{R}$)

Ce fait découle immédiatement d'un corollaire donné par Rockafellar (Cfr [8], corollaire 8.7.1 page 70.) dont l'énoncé est le suivant :

"Soit f une fonction convexe propre et s.c.i.
Si $\{x \mid f(x) \leq \alpha\}$ est non vide et borné pour un α, il est borné pour tout α."

(f est une fonction définie sur un espace vectoriel topologique X de dimension finie et à valeurs dans $\overline{\mathbb{R}}$)

Dans notre cas nous avons :

- $-c(\cdot, \bar{\bar{v}})$ est convexe, propre (c est à valeurs dans $\mathbb{R}$) et s.c.i (en fait elle est même continue car c'est une fonction convexe définie sur un espace de dimension finie et à valeurs dans $\mathbb{R}$).
- $\{u \mid -c(u, \bar{\bar{v}}) \leq 0\}$ est non vide et borné (cfr compact)

En appliquant le corollaire nous obtenons :

$\{u \mid c(u, \bar{\bar{v}}) \geq \lambda\}$ est borné $\quad \forall \lambda \in \mathbb{R}$.

En outre, comme $c(\cdot, \bar{\bar{v}})$ est continue donc s.c.s

$\{u \mid c(u, \bar{\bar{v}}) \geq \lambda\}$ est fermé $\quad \forall \lambda \in \mathbb{R}$.

Comme nous travaillons dans un espace de dimension finie nous obtenons :

$\forall \lambda \in \mathbb{R} \qquad \{u \mid c(u, \bar{\bar{v}}) \geq \lambda\}$ est compact.

Voyons que ceci implique que le suprémum est atteint c'est à dire qu'il existe un élément u de U tel que :

$$\bar{c}(G(\tilde{x}), \bar{\bar{v}}) = c(u, \bar{\bar{v}}) \quad \text{et} \quad G(\tilde{x}) \leq u$$

En effet, $\bar{c}(G(\tilde{x}), \bar{\bar{v}}) \overset{def}{=} \operatorname{Sup}\{c(u', \bar{\bar{v}}) \mid u' \in U,\ G(\tilde{x}) \leq u'\}$

$= \operatorname{Sup}\{c(u', \bar{\bar{v}}) \mid u' \in U;\ G(\tilde{x}) \leq u';\ c(u', \bar{\bar{v}}) \geq c(G(\tilde{x}), \bar{\bar{v}})\}$

Or, $\{u' \in U \mid G(\tilde{x}) \leq u'$ et $c(u', \bar{v}) \geq c(G(\tilde{x}), \bar{v})\}$ est compact car intersection du fermé $\{u' \in U \mid G(\tilde{x}) \leq u'\}$ et du compact $\{u' \in U \mid c(u', \bar{v}) \geq c(G(\tilde{x}), \bar{v})\}$ (cf partie 1 de la démonstration).

Par conséquent, nous prenons le supremum d'une fonction continue sur un compact non vide (cf $G(\tilde{x})$ y appartient) il est donc atteint.

2) Un exemple simple montre également que l'on ne peut affaiblir les hypothèses. En ce sens que $\bar{v} \in c\text{-}\partial h(\sigma)$ n'est suffisant pour qu'un maximisant $\tilde{x}$ de $l(\cdot, \bar{v})$ soit une solution optimale de (P).

En effet, posons $U = \mathbb{R}$, $V = \mathbb{R}_+$, $c(u, v) = -v|u|$, $u \in U, v \in \mathbb{R}_+$ et considérons le programme mathématique défini par :

$$\begin{cases} f_0(x) = -|x| & x \in X = \mathbb{R} \\ G(x) = x \in U & \\ C = \{\sigma\} & \text{où } \sigma = 0. \end{cases}$$

Calculons la fonction de perturbation h.

$$\begin{aligned} h(u) &\overset{\text{déf}}{=} \inf\{f_0(x) \mid x \in X, G(x) \leq u\} && \forall u \in U \\ &= \inf\{-|x| \mid x \in X, x - u = 0\} && \forall u \in U \\ &= -|u| && \forall u \in U \end{aligned}$$

Remarquons que le réel 1 appartient à $c\text{-}\partial h(0)$ car

$$h(u) - h(0) \geq c(u, 1) - c(0, 1) \qquad \forall u \in U$$

(cfr définition de h et de c).

Calculons le lagrangien :

$$l(x,v) \overset{\text{déf}}{=} -f_0(x) + \bar{c}(G(x),v) \qquad x \in X,\ v \in V$$

où $\bar{c}(u,v) = \operatorname{Sup}\{c(u',v) \mid u' \in U,\ u \leq u'\}$

Par conséquent,

$$\bar{c}(G(x),v) = \operatorname{Sup}\{-v|u'| \mid u' \in U,\ u' - x = 0\}$$
$$= -v|x|$$

donc, $l(x,v) = (1-v)|x| \qquad \forall x \in X,\ \forall v \in V$

et $l(x,1) = 0$

Tout réel x maximise $l(x,1)$. Cependant, la solution optimale unique du problème initial est zéro.

III.4. Application d'une méthode de pénalisation exacte

Dans ce paragraphe, nous allons considérer la question de l'existence de c-sous gradients de la fonction de perturbation en σ.

Nous avons vu dans le théorème III.3 que ces c-sous gradients permettent de construire des multiplicateurs exacts par le biais de la propriété de flexibilité pour la fonctionnelle couplante.

Définition III.2

Considérons une fonctionnelle a de $\overline{\mathbb{R}}^U$ et deux éléments u_o, u de U. Nous définissons la dérivée directionnelle inférieure (supérieure) de a en u_o dans la direction u, notée $\underline{D}a(u_o;u)$ $(\overline{D}a(u_o;u))$, par :

$$\underline{D}\,a(u_o,u) = \liminf_{\substack{\lambda \to 0 \\ >}} \left\{ \frac{a(u_o+\lambda u) - a(u_o)}{\lambda} \right\}$$

$$\overline{D}\,a(u_o,u) = \limsup_{\substack{\lambda \to 0 \\ >}} \left\{ \frac{a(u_o+\lambda u) - a(u_o)}{\lambda} \right\}$$

Définition III.3

Une fonctionnelle a de $\overline{\mathbb{R}}^U$ est dite "LOCALEMENT CONVEXE" en un élément u_o de U si il existe un voisinage N de u_o tel que :

$\forall u \in N \ , \ \forall \lambda \in]0,1[$

$$a(\lambda u + (1-\lambda)u_o) \leq \lambda a(u) + (1-\lambda)\, a(u_o)$$

Nous dirons qu'une fonctionnelle a appartenant à $\overline{\mathbb{R}}^V$ est "LOCALEMENT CONCAVE" en un élément u_0 de V si $-a$ est localement convexe en u_0.

La proposition suivante fournit des conditions suffisantes pour la c-sous différentiabilité de h dans le cas où les propriétés du premier ordre de h donnent à elles seules une information suffisante.

Proposition III.3

Supposons que la fonctionnelle couplante c soit flexible en σ. Si a, appartenant à $\overline{\mathbb{R}}^V$, est localement convexe, finie en σ, c-tempérée et s'il existe un élément v de V et un voisinage N de σ tels que:

$$\begin{cases} c(\cdot, v) \text{ soit localement concave en } \sigma \\ \forall u \in N \quad \underline{D}\, a(\sigma, u) \geq \bar{D}\, c_v(\sigma, u) \end{cases}$$

alors c-$\partial a(\sigma)$ est non vide.

($\bar{D} c_v(\sigma, u)$ désigne la dérivée directionnelle de $c(\cdot, v)$)

Démonstration

Dans une première étape nous démontrons que: a localement convexe en σ implique l'existence d'un voisinage $\tilde{N}$ de σ tel que:

$$\forall u \in \tilde{N} \quad \underline{D}\, a(\sigma, u) \leq a(u) - a(\sigma)$$

Par définition de la convexité locale de a en σ, il existe un voisinage $\bar{N}$ de σ tel que :

$$a(\lambda u + (1-\lambda)\sigma) \leq \lambda a(u) + (1-\lambda)\, a(\sigma) \qquad \forall u \in \bar{N},\ \forall \lambda \in]0,1[$$

Comme σ est le neutre pour l'addition dans U, nous obtenons :

$$\frac{a(\lambda u) - a(\sigma)}{\lambda} \leq a(u) - a(\sigma) \qquad \forall u \in \bar{N},\ \forall \lambda \in]0,1[$$

Par conséquent, en prenant la limite inférieure du premier membre de l'inégalité nous avons :

$$\sup_{\varepsilon \geq 0}\ \inf_{\lambda \in]0,\varepsilon[}\ \frac{a(\lambda u) - a(\sigma)}{\lambda} = \underline{D}\, a(\sigma, u) \leq a(u) - a(\sigma) \quad , \ \forall u \in \bar{N}$$

Nous allons maintenant prouver l'existence d'un voisinage N' de σ tel que :

$$\forall u \in N' \quad a(u) - a(\sigma) \geq c(u, v) - c(\sigma, v)$$

Par hypothèse, $c(., v)$ est localement concave en σ, donc il existe un voisinage $\bar{\bar{N}}$ de σ tel que :

$$c_v(\lambda u + (1-\lambda)\sigma) \geq \lambda c_v(u) + (1-\lambda)\, c_v(\sigma) \qquad \forall u \in \bar{\bar{N}},\ \forall \lambda \in]0,1[$$

En procédant de façon similaire à ci-dessus, nous obtenons :

$$\inf_{\varepsilon \geq 0}\ \sup_{\lambda \in]0,\varepsilon[}\ \frac{c_v(\lambda u) - c_v(\sigma)}{\lambda} = \bar{D} c_v(\sigma, u) \geq c_v(u) - c_v(\sigma) \quad , \ \forall u \in \bar{\bar{N}}$$

Posons $N' = N \cap \bar{N} \cap \bar{\bar{N}}$, N' est un voisinage de σ et

$$a(u) - a(\sigma) \underset{(u \in \bar{N})}{\geq} \underline{D}\, a(\sigma, u) \underset{(u \in N)}{\geq} \bar{D} c_v(\sigma, u) \underset{(u \in \bar{\bar{N}})}{\geq} c_v(u) - c_v(\sigma) \qquad \forall u \in N'$$

Nous allons enfin pouvoir conclure que le c-sous-différentiel de a en σ est non vide.

En effet, par hypothèse a est c-tempérée donc il existe un élément $\bar{v}$ de V et un réel $\bar{\eta}$ tels que :

$$c(u,\bar{v}) + \bar{\eta} \leq a(u) \qquad \forall u \in U$$

Par conséquent,

$$c(u,\bar{v}) + \bar{\eta} - a(\sigma) \leq a(u) - a(\sigma) \qquad \forall u \in U \qquad (1)$$

Posons $\eta = \bar{\eta} - a(\sigma)$. η est fini car $\bar{\eta}$ et $a(\sigma)$ le sont.
Appliquons maintenant la flexibilité de c en σ en prenant : $N = N'$, η, $v' = \bar{v}$, $v'' = v$

Il existe donc un voisinage N'' de σ tel que N'' est inclus dans N' et un élément $\bar{\bar{v}}$ de V vérifiant :

$$\left| \begin{array}{ll} c(u,\bar{\bar{v}}) \leq c(u,\bar{v}) + \eta \underset{(1)}{\leq} a(u) - a(\sigma) & \forall u \notin N'' \quad (2) \\ \\ c(u,\bar{\bar{v}}) \leq c(u,v) \leq a(u) - a(\sigma) & \forall u \in N'' \quad (3) \end{array} \right.$$

où la seconde inégalité découle de la première étape de démonstration et de la normalisation de c en σ.

Donc, en résumant (2) et (3) en une seule inégalité nous obtenons :

$$c(u,\bar{\bar{v}}) - c(\sigma,\bar{\bar{v}}) \leq a(u) - a(\sigma) \qquad \forall u \in U$$

En outre, $a(\sigma)$ est fini.

Par conséquent, $\bar{\bar{v}}$ appartient au c-sous-différentiel de a en σ.

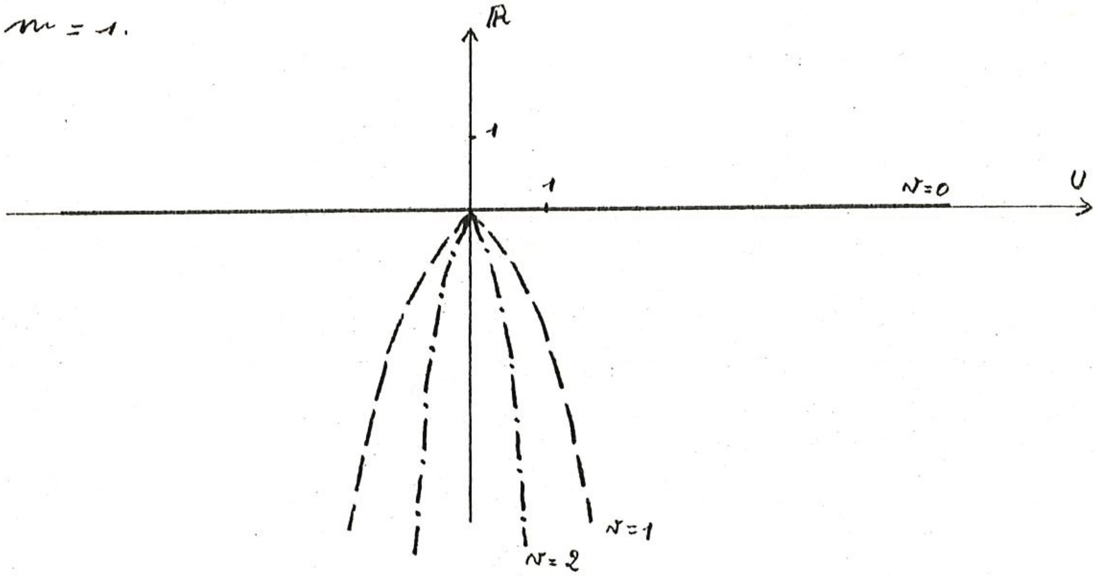
Nous allons maintenant étudier un exemple pour lequel, en appliquant la dualité non convexe, nous obtenons des résultats d'existence de multiplicateurs exacts.

<u>Exemple</u>

Posons $U = \mathbb{R}^m$, $V = \mathbb{R}^m_+$ et

$$c(u,v) = -\sum_{i=1}^{m} (e^{v_i |u_i|} - 1) \qquad u = (u_1, u_2 \cdots u_m) \in \mathbb{R}^m,\ v = (v_1, \cdots v_m) \in \mathbb{R}^m_+$$

Représentons $c(\cdot, v)$ pour différentes valeurs de v dans le cas où $m = 1$.



Considérons le programme mathématique défini par :

$$\left| \begin{array}{l} G = (g_1, g_2 \cdots g_m) \quad \text{où } g_i : X \to \mathbb{R} \quad i = 1, \dots m \\ C = \mathbb{R}^m_- \quad \text{le cône négatif dans } U \end{array} \right.$$

Nous supposons, en outre, que :

1) $f_0, g_1 \cdots g_m$ sont des fonctionnelles convexes sur l'ensemble convexe X.
2) f_0 est bornée inférieurement sur X.
3) La condition de Slater est satisfaite c'est à dire qu'il existe un élément $\bar{x}$ de X tel que $G(\bar{x})$ appartient à l'intérieur de C. Ceci est équivalent à : $\quad g_i(\bar{x}) < 0 \quad$ pour $i = 1, 2 \dots m$

Ces hypothèses impliquent que h est convexe sur U, finie et continue en σ. (cfr chapitre I, paragraphe I.3)

Nous pouvons donc employer le résultat d'analyse fonctionnelle convexe suivant :

> Si f est une fonctionnelle convexe sur X, finie et continue en un point x_0 de X alors le sous différentiel de f en x_0 est non vide.

(démonstration voir [2] page 353)

Par conséquent, si nous considérons le produit scalaire sur $\mathbb{R}^m$ comme étant le crochet de dualité, nous pouvons appliquer ce théorème à h et nous obtenons :

$$\exists \bar{u} \in \mathbb{R}^m \text{ tel que } \quad h(u) - h(\sigma) \geq \langle u, \bar{u} \rangle \quad \forall u \in \mathbb{R}^m \qquad (I)$$

Nous allons maintenant vérifier que nous sommes bien dans les conditions d'application de la proposition III.3

1) <u>En premier lieu nous prouvons que :</u>

> $\forall u \in \mathbb{R}^m$ tel que $|u| < 1$ nous avons :
>
> $$\underline{D} h(\sigma, u) \geq \langle u, \bar{u} \rangle \geq - \sum_{i=1}^{m} |\bar{u}_i| . |u_i| \geq \bar{D} c_v(\sigma, u) \qquad (II)$$
>
> où v est un élément de V vérifiant $v_i \geq |\bar{u}_i|$ $i = 1, 2 .. m$

Démonstration

Par une démonstration tout à fait similaire à celle de la première étape de la proposition III.3, nous obtenons :

$$\underline{D}\, h(\sigma, u) \geq h(u) - h(\sigma) \qquad \forall u \in U$$

Donc, par (I) et par définition du produit scalaire, nous avons :

$$\underline{D}\, h(\sigma, u) \geq \langle u, \bar{u} \rangle \geq -\sum_{i=1}^{m} |\bar{u}_i| . |u_i| \qquad \forall u \in U$$

$$\text{Or, } \bar{D} c_v(\sigma, u) \overset{\text{déf}}{=} \limsup_{\lambda \underset{>}{\to} 0} \frac{c(\sigma + \lambda u, v) - c(\sigma, v)}{\lambda}$$

$$= \limsup_{\lambda \underset{>}{\to} 0} \frac{-\sum_{i=1}^{m} (e^{\lambda v_i |u_i|} - 1)}{\lambda}$$

$$= -\sum_{i=1}^{m} v_i |u_i|$$

Par conséquent, si v_i est supérieur ou égal à $|\bar{u}_i|$ pour $i = 1 \ldots m$ l'inégalité (II) est bien vérifiée.

2) Vérifions que c_v est concave pour tout v dans V

Démonstration

Considérons v un élément arbitraire de V.

Soient u, u' appartenant à U et un réel λ quelconque compris entre zéro et un.

$$c_v(\lambda u + (1-\lambda) u') = -\sum_{i=1}^{m} \left(e^{v_i |\lambda u_i + (1-\lambda) u'_i|} - 1\right)$$

$$= -\sum_{i=1}^{m} \left(e^{\lambda v_i |u_i| + (1-\lambda) v_i |u'_i|} - 1\right)$$

En utilisant la convexité de l'exponentielle dans $\mathbb{R}$, on obtient :

$$c_v(\lambda u + (1-\lambda) u') \geq -\sum_{i=1}^{m} \left(\lambda e^{v_i |u_i|} + (1-\lambda) e^{v_i |u'_i|} - \lambda - (1-\lambda)\right)$$

$$\geq \lambda c_v(u) + (1-\lambda) c_v(u')$$

3) Voyons que C est flexible en σ

Démonstration

Nous devons prouver que pour tout voisinage N de σ, pour tout v', v'' éléments de V et tout réel η, il existe un voisinage de σ N' inclus dans N et un élément v de V tels que :

$$-\sum_{i=1}^{m}(e^{v_i|u_i|}-1) \leq -\sum_{i=1}^{m}(e^{v_i'|u_i|}-1)+\eta \qquad \forall u \notin N' \qquad (1)$$

$$-\sum_{i=1}^{m}(e^{v_i|u_i|}-1) \leq -\sum_{i=1}^{m}(e^{v_i''|u_i|}-1) \qquad \forall u \in N' \qquad (2)$$

. Prenons pour N' une boule $B(\sigma,r)$ où

$$B(\sigma,r) \overset{déf}{=} \{u \in U \mid \max_{i\in\{1,..m\}} |u_i| \leq r\}$$

et r est choisi de telle sorte que $B(\sigma,r)$ soit contenue dans N

. Considérons un élément v de V tel que :

a) $v_i \geqslant v_i'' \qquad i=1,..m$

(ceci implique immédiatement que l'inégalité (2) est vérifiée)

b) $v_i \geqslant v_i' - \eta \cdot \dfrac{1}{[r+r^2(\inf_i v_i')].m} \qquad i=1,..m$

c) $v_i \geqslant v_i' \quad i=1,..m$

Voyons que si v_i vérifie b) et c) pour $i=1,..m$, l'inégalité (1) est satisfaite.

Exprimons l'inégalité (1) de manière différente, soit :

$$\sum_{i=1}^{m}(e^{v_i|u_i|}-e^{v_i'|u_i|}) \geqslant -\eta \qquad \forall u \notin N'$$

Développons le premier membre de l'inégalité précédente :

$$\sum_{i=1}^{m} \left(e^{v_i|u_i|} - e^{v_i'|u_i|}\right) = \sum_{i=1}^{m} e^{v_i'|u_i|}\left(e^{(v_i - v_i')|u_i|} - 1\right)$$

$$\geq \sum_{i=1}^{m} (1 + v_i'|u_i|)\,((v_i - v_i')\,|u_i|)$$

$$\geq \sum_{i=1}^{m} (v_i - v_i')|u_i| + \sum_{i=1}^{m} (v_i - v_i')\,v_i'\,|u_i|^2$$

Nous avons supposé $v_i \geq v_i'$ $i = 1 \ldots m$ pour réaliser ce développement et comme nous ne devons considérer que les éléments u n'appartenant pas à N', nous avons l'inégalité suivante :

$$\sum_{i=1}^{m} \left(e^{v_i|u_i|} - e^{v_i'|u_i|}\right) \geq \inf_{i \in \{1,\ldots m\}} (v_i - v_i')\, r.m + \inf_{i \in \{1,\ldots m\}} (v_i - v_i') . \inf_{i \in \{1,\ldots m\}} v_i' . r^2 . m$$

Pour que l'inégalité (1) soit vérifiée il suffit donc que :

$$\inf_{i \in \{1 \ldots m\}} (v_i - v_i')\left[r + \inf_{i \in \{1 \ldots m\}} v_i' . r^2\right] \geq \frac{-\eta}{m}$$

Ce qui revient à la condition (2).

Donc l'élément v de V qui convient doit satisfaire.

$$v_i \geq \max\left(v_i', v_i'', v_i' - \eta . \frac{1}{[r + r^2 \inf_i v_i']\, m}\right) \quad \forall i \in \{1, 2, \ldots m\}$$

4) <u>Nous devons encore prouver que h est c-tempérée</u>

<u>Démonstration</u>

Par hypothèse, f_0 est bornée inférieurement sur X, cela implique :
$\exists M \in \mathbb{R}$ tel que $f_0(x) \geq M$ $\forall x \in X$

Or, $\forall u \in U \quad h(u) = \inf_{x \in X} \phi(x,u) \geq \inf_{x \in X} f(x) \geq M$

Donc, si on choisit v tel que : $v_i = 0 \quad \forall i \in \{1, \dots m\}$

nous obtenons :

$$h(u) \geq c(u,v) + M \quad \forall u \in U$$

Par conséquent, h est bien c-tempérée.

Les hypothèses de la proposition III.3 sont donc bien remplies et nous pouvons conclure que le c-sous-différentiel de h en σ est différent du vide. c'est à dire :

$$\exists \bar{v} \in V \quad tq \quad \bar{v} \in c\text{-}\partial h(\sigma)$$

ou

$$h(u) - h(\sigma) \geq c(u,\bar{v}) - c(\sigma,\bar{v}) \quad \forall u \in U$$

Considérons maintenant $\bar{\bar{v}}$ appartenant à V construit grâce à la propriété de flexibilité de c en σ à partir d'un voisinage N de σ, un réel η strictement négatif et pour $v' = v'' = \bar{v}$.

donc : $\bar{\bar{v}}_i \geq \max\left(\bar{v}_i, \bar{v}_i - \eta \cdot \dfrac{1}{[r + r^2 \inf_i \bar{v}_i] \cdot m}\right) \quad i = 1 \dots m$

où r est un réel positif déterminé par N

Comme η est strictement négatif il s'ensuit :

$$\bar{\bar{v}}_i > \bar{v}_i \qquad i = 1 \dots m$$

Par conséquent, grâce à la croissance de l'exponentielle nous avons :

$$c(u,\bar{\bar{v}}) = -\left(\sum_{i=1}^{m} e^{\bar{\bar{v}}_i |u_i|} - 1\right) < -\left(\sum_{i=1}^{m} e^{\bar{v}_i |u_i|} - 1\right) = c(u,\bar{v}) \quad \forall u \in N \setminus \{0\}$$

En outre, $c(\cdot,\bar{\bar{v}})$ est concave, U est un espace vectoriel de dimension finie et $\{u \mid c(u,\bar{\bar{v}}) \geq 0\} = \{0\}$ est compact.

En appliquant la remarque 1. sur le théorème III.3 nous obtenons que $\sup\{c(u,\bar{\bar{v}}) \mid u \in U,\ G(\tilde{x}) \leq u\}$ est atteint.

Par conséquent, nous vérifions toutes les hypothèses du théorème III.3 et donc, tout maximisant $\bar{\bar{x}}$ de $\ell(\cdot,\bar{\bar{v}})$ sera une solution optimale du problème initial.

ANNEXE . I

1. Lemme donnant des conditions suffisantes pour qu'une fonctionnelle de $\mathbb{R}^{U\times V}$ soit de type aiguille en un point de U.

Enoncé :

Supposons que la topologie sur U soit engendrée par une métrique d.

Soient u_0 appartenant à U et $s : \mathbb{R}_+ \longrightarrow \mathbb{R}_+$ une fonction strictement croissante telle que :
$$\begin{cases} s(0) = 0 \\ \exists K \in \mathbb{R} \ : \forall x \in \mathbb{R}_+ \ \ s(2x) \leq K s(x) \end{cases}$$

Supposons que la fonctionnelle couplante c de $\mathbb{R}^{U\times V}$ a la propriété suivante :

$\forall v \in V$, $\exists u_1 \in U$, $\eta_1 \in \mathbb{R}$, $\rho_1 \in \mathbb{R}$ tels que
$$c(\cdot, v) \geq \rho_1 \, s[d(\cdot, u_1)] + \eta_1$$
et que la collection des c-fonctionnelles élémentaires finies comprend toutes les fonctionnelles $-\rho \, s[d(\cdot, u_0)]$ où ρ est un réel positif.

Alors c est de type aiguille en u_0.

Remarque

Tous les monômes de $\mathbb{R}_+$ satisfont les conditions imposées sur s dans l'énoncé du lemme.

exemple : $s : \mathbb{R}_+ \longrightarrow \mathbb{R}_+ : x \longmapsto 2x^2$

. $s(0) = 0$

. $s(2x) = 2.4x^2 \leq 4\, s(x)$

Démonstration du lemme:

Soient N un voisinage arbitraire de u_0, v un élément de V et η un réel quelconque.

Nous devons démontrer l'existence d'un voisinage de u_0, N' inclus dans N et d'un élément v' de V tels que:

$$\begin{cases} c(u,v') - c(u_0,v') \leq c(u,v) + \eta & \forall u \notin N' \quad (1) \\ c(u,v') - c(u_0,v') \leq 0 & \forall u \in N' \quad (2) \end{cases}$$

Considérons une boule ouverte de centre u_0 et de rayon $\delta > 0$ incluse dans N et appelons-la N'.
Par hypothèse, pour l'élément v choisi dans V, il existe un élément u_1 de U et deux réels η_1, ρ_1 tels que:

$$c(u,v) \geq \rho_1 s[d(u,u_1)] + \eta_1 \qquad \forall u \in U$$

Posons $\rho_2 = |\rho_1|$ et $\eta_2 = \eta_1 + \eta$, nous avons alors:

$$c(u,v) + \eta \geq -\rho_2 s[d(u,u_1)] + \eta_2 \qquad \forall u \in U \qquad (3)$$

Dans une première étape nous allons vérifier que:

(I) Si ρ est suffisamment grand nous avons:
$$-\rho s[d(u,u_0)] \leq -\rho_2 s[d(u,u_1)] + \eta_2 \qquad \forall u \notin N'$$

Cela nous permettra d'obtenir la partie (1) de la thèse.

En effet, (I) implique, en utilisant l'inégalité (3), que:
$$\exists \rho \in \mathbb{R}, \rho \geq 0 \ \text{tq:} \ -\rho s[d(u,u_0)] \leq c(u,v) + \eta \qquad \forall u \notin N'$$

Par hypothèse, nous savons que $-\rho\, s[d(u,u_0)]$ est une c-fonctionnelle élémentaire finie c'est à dire:

$\exists v' \in V,\ \exists \gamma' \in \mathbb{R}$ tel que $-\rho\, s[d(u,u_0)] = c(u,v') + \gamma' \quad \forall u \in U$

Or, comme $s(0)=0$, nous avons: $\gamma' = -c(u_0, v')$

donc,

$\exists v' \in V$ tel que : $c(u,v') - c(u_0,v') \leq c(u,v) + \gamma \quad \forall u \notin N'$

Par conséquent si (I) est satisfait la partie (1) est prouvée.

Vérifions (I):

Pour cela nous allons envisager deux cas:

1) Nous considérons les éléments u de U n'appartenant pas à N' (i.e $d(u,u_0) \geq \delta$) et qui vérifient l'inégalité suivante:

$$d(u,u_0) \leq d(u_0,u_1) \qquad (4)$$

Dans ce cas, en utilisant l'hypothèse de croissance de s et l'inégalité triangulaire pour d, nous obtenons:

$$s[d(u,u_1)] \leq s[d(u,u_0) + d(u_0,u_1)]$$

donc, en employant (4) nous avons:

$$s[d(u,u_1)] \leq s[2\,d(u_0,u_1)]$$

Ce qui implique, par hypothèse :

$$s[d(u,u_1)] \leq K\, s[d(u_0,u_1)]$$

Puisque ρ_2 est positif et u n'appartient pas à N', nous déduisons:

$$\forall \rho \geq 0 \quad -\rho_2\, s[d(u,u_1)] + \rho\, s[d(u,u_0)] \geq -\rho_2 K\, s[d(u_0,u_1)] + \rho\, s[\delta]$$

donc,

$$\forall \rho \geq 0 \quad -\rho_2\, s[d(u,u_1)] + \rho_2 K\, s[d(u_0,u_1)] - \rho\, s[\delta] \geq -\rho\, s[d(u,u_0)]$$

Pour avoir (I), il suffit de vérifier que le premier membre de cette inégalité est inférieur $-\rho_2 s[d(u,u_1)] + \eta_2$ pour ρ suffisamment grand.

Cela revient donc à prouver l'existence d'un réel ρ positif tel que.

$$\rho_2 K s[d(u_0,u_1)] - \eta_2 \leq \rho s[\delta]$$

C'est immédiat pour ρ suffisamment grand puisque le premier membre de l'inégalité est fixé et $s[\delta]$ est strictement positif.

2) Nous considérons toujours les éléments u de U n'appartenant pas à N' (i.e $d(u,u_0) \geq \delta$) mais qui vérifient, en outre, l'inégalité $d(u,u_0) > d(u_0,u_1)$

En procédant de manière similaire au cas 1) nous obtenons:

$$s[d(u,u_1)] \leq K s[d(u_0,u)]$$

Et finalement,

$$\forall \rho \in \mathbb{R} \quad -\rho_2 s[d(u,u_1)] + \rho s[d(u,u_0)] \geq (-\rho_2 K + \rho) s[d(u,u_0)]$$

Donc, puisque u n'appartient pas à N' et s est croissante nous avons:

$$\forall \rho \geq \rho_2 K \quad -\rho_2 s[d(u,u_1)] + \rho s[d(u,u_0)] \geqslant (-\rho_2 K + \rho) s[\delta]$$

Par conséquent,

$$\forall \rho \geq \rho_2 K \quad -\rho_2 s[d(u,u_1)] - (\rho - \rho_2 K) s[\delta] \geqslant -\rho s[d(u,u_0)]$$

En raisonnant de la même façon que dans le cas 1) nous voyons que (I) est vérifié.

La partie (1) de la thèse est donc démontrée.

Il nous reste encore à vérifier (2) mais ceci est immédiat car:

$$c(u, v') - c(u_0, v') = -\rho s[d(u, u_0)] \qquad \forall u \in U$$

or, $\rho \geq 0$ et $s[d(u, u_0)] \geq 0$

donc,

$$c(u, v') - c(u_0, v') \leq 0 \qquad \forall u \in N'$$

2. Vérification des exemples.

Exemple 1.a : $U = \mathbb{R}^m$, $V = \mathbb{R}_+^m$

$$c(u, v) = -\sum_{i=1}^{m} v_i |u_i| \quad , u = (u_1, .. u_m) \in \mathbb{R}^m, v = (v_1, .. v_m) \in \mathbb{R}_+^m$$

Nous devons prouver que c est de type aiguille à l'origine de $\mathbb{R}^m$.

Pour cela, nous allons démontrer que c vérifie bien toutes les hypothèses du lemme.

a) U est muni d'une topologie engendrée par une métrique d définie par:

$$d(x, y) = \sum_{i=1}^{m} |x_i - y_i| \qquad \forall x, y \in \mathbb{R}^m.$$

b) Considérons la fonction $s : \mathbb{R}_+ \longrightarrow \mathbb{R}_+ : x \longmapsto s(x) \overset{\text{déf}}{=} x$

De façon immédiate, nous voyons que s convient car:

. s est croissante

. $s(2x) \leq 2\, s(x) \quad \forall x \in \mathbb{R}_+$

. $s(0) = 0$

c) Nous devons démontrer que :

$\forall v \in V, \exists u_1 \in U, \eta_1 \in \mathbb{R}, \rho_1 \in \mathbb{R}$ tels que :

$$c(u,v) \geq \rho_1 \, s[d(u,u_1)] + \eta_1 \qquad \forall u \in U$$

c'est à dire,

$$-\sum_{i=1}^{m} v_i |u_i| \geq \rho_1 \sum_{i=1}^{m} |u_{1_i} - u_i| + \eta_1 \qquad \forall u \in U \qquad (1)$$

Soit v un élément arbitraire dans V

Choisissons
$$\left| \begin{array}{l} u_1 = (0,0,\dots 0) \in \mathbb{R}^m \\ \rho_1 = -\max\limits_{i \in \{1,2,\dots m\}} (v_i) \\ \\ \eta_1 = 0 \end{array} \right.$$

Nous voyons que l'inégalité (1) est vérifiée pour ce choix des éléments u_1, ρ_1, η_1

d) Nous devons encore établir l'assertion suivante :

$\forall \rho \geq 0, \exists v \in V, \eta \in \mathbb{R}$ tels que :

$$-\rho \, s[d(u,u_0)] = c(u,v) + \eta \qquad \forall u \in U.$$

(u_0 désignant l'origine de $\mathbb{R}^m$)

Choisissons
$$\left| \begin{array}{l} v = (v_1, \dots v_m) \text{ tel que : } v_i = \rho \quad \forall i \in \{1,2\dots m\} \\ \eta = 0 \end{array} \right.$$

Nous savons, par définition de s et de d, que :

$$-\rho \, s[d(u,u_0)] = -\rho \sum_{i=1}^{m} |u_i|$$

Par conséquent, $-\rho \, s[d(u,u_0)] = c(u,v) + \eta$

Nous pouvons maintenant appliquer le lemme et nous avons donc que c est de type aiguille en l'origine de $\mathbb{R}^m$.

Exemple 2. a :

$U = H$ un espace de Hilbert , $V = \mathbb{R}_+ \times H$

$c(u,v) = - v_0 |u|^2 - \langle w, u \rangle$; $v = (v_0, w) \in \mathbb{R}_+ \times H$, $u \in H$

($\langle .,. \rangle$ désigne le produit scalaire dans H et $|.|$ la norme induite.)

Ici, nous allons démontrer que le lemme est applicable à c en tout élément u_0 de U . Donc, c est de type aiguille en tout élément de U.

Vérifions que c satisfait les hypothèses du lemme pour un élément u_0 arbitraire dans U.

a) U est muni d'une topologie engendrée par la métrique d déduite du produit scalaire et définie par :

$$d(x,y) = \sqrt{\langle x-y, x-y \rangle} \qquad \forall x, y \in H.$$

b) Considérons la fonction $s : \mathbb{R}_+ \longrightarrow \mathbb{R}_+ : x \longmapsto s(x) \overset{\text{déf}}{=} x^2$

s convient car :

. s est croissante sur $\mathbb{R}_+$

. $s(2x) \leq 4 . s(x)$

. $s(0) = 0$

c) Soit $v = (v_0, w)$ arbitraire dans V, nous devons vérifier que :

$\exists u_1 \in V$, $\eta_1 \in \mathbb{R}$, $\rho_1 \in \mathbb{R}$ tels que :

$$c(u,v) \geq \rho_1 \, s[d(u,u_1)] + \eta_1 \qquad \forall u \in U$$

c'est à dire,

$$- v_0 |u|^2 - \langle w, u \rangle \geq \rho_1 \langle u - u_1, u - u_1 \rangle + \eta_1 \qquad \forall u \in U$$

En transformant la dernière inégalité grâce aux propriétés du produit scalaire, nous en déduisons une formulation équivalente :

$$(-\nu_0 - \rho_1)\,|u|^2 - \langle w - 2\rho_1 u_1 , u \rangle - \rho_1 |u_1|^2 - \eta_1 \geq 0 \qquad \forall u \in U \quad (1)$$

(1) est vérifié si nous choisissons

$$\left|\begin{array}{l} \rho_1 \leq -\nu_0 \quad \text{et } \rho_1 \neq 0 \\ u_1 = \dfrac{w}{2\rho_1} \\ \\ \eta_1 = -\rho_1 |u_1|^2 \end{array}\right.$$

d) Il nous reste à démontrer la proposition suivante :

$\forall \rho \geq 0$, $\exists \nu \in V$, $\eta \in \mathbb{R}$ tels que :

$$-\rho\, s[d(u,u_0)] = c(u,\nu) + \eta \qquad \forall u \in U$$

(u_0 désignant l'élément arbitraire de U fixé au départ)

Choisissons $\left|\begin{array}{ll} \nu = (\rho, -2\rho u_0) & \in V \\ \eta = -\rho |u_0|^2 & \in \mathbb{R} \end{array}\right.$

Nous avons alors :

$$\begin{aligned} -\rho\, s[d(u,u_0)] &= -\rho \cdot \langle u - u_0 , u - u_0 \rangle && \forall u \in U \\ &= -\rho |u|^2 - \rho |u_0|^2 + 2\rho \langle u, u_0 \rangle && \forall u \in U \\ &= c(u,\nu) + \eta \end{aligned}$$

Par conséquent, toutes les hypothèses du lemme sont vérifiées en un élément u_0 arbitraire dans U. Il s'ensuit que c est de type aiguille en tout élément de U.

ANNEXE II

Démonstration du théorème II.1

Dans une première étape, nous allons démontrer que si c est une fonctionnelle de $\mathbb{R}^{U\times V}$ pointue en un élément u_0 de U, alors toute fonctionnelle a de $\overline{\mathbb{R}}^U$ c-tempérée et s.c.i en u_0 vérifie : $a(u_0) = a^{cc}(u_0)$.

Considérons donc c appartenant à $\mathbb{R}^{U\times V}$ pointue en u_0 et une fonctionnelle arbitraire a de $\overline{\mathbb{R}}^U$ c-tempérée et s.c.i en u_0.

Nous envisageons deux cas :

a) $a(u_0)$ fini

Soit ε, un réel strictement positif arbitraire.

Puisque a est s.c.i en u_0, il existe un voisinage N de u_0 tel que :

$$a(u) \geq a(u_0) - \varepsilon \qquad \forall u \in N \qquad (1)$$

Par hypothèse, a est c-tempérée, donc il existe un élément v de V et un réel η tels que :

$$a(u) \geq c(u,v) + \eta \qquad \forall u \in U \qquad (2)$$

En utilisant le caractère pointu de c en u_0, il existe v' appartenant à V et un voisinage N' de u_0, N' inclus dans N, tels que.

$$\begin{cases} c(u,v') - c(u_0,v') \leq c(u,v) + \eta - a(u_0) + \varepsilon & \forall u \notin N' \quad (3) \\ c(u,v') - c(u_0,v') \leq \varepsilon & \forall u \in N' \quad (4) \end{cases}$$

Si nous transformons (3) en utilisant (2) et si nous additionnons membre à membre les inégalités (11 et (4), nous obtenons :

$$\left| \begin{array}{ll} c(u,v') - c(u_0,v') + a(u_0) - \varepsilon \leq a(u) & \forall u \notin N' \\ c(u,v') - c(u_0,v') + a(u_0) - 2\varepsilon \leq a(u) & \forall u \in N' \end{array} \right.$$

Donc,

$$c(u,v') - c(u_0,v') + a(u_0) - 2\varepsilon \leq a(u) \quad \forall u \in U$$

Nous obtenons une c-minorante affine de a et par conséquent, de a^{cc}. (cfr propriétés des fonctionnelles c-conjuguées)

En prenant $u = u_0$, nous déduisons :

$$a(u_0) - 2\varepsilon \leq a^{cc}(u_0)$$

Cette inégalité est valable pour tout réel ε strictement positif. Par conséquent,

$$a(u_0) \leq a^{cc}(u_0)$$

Par la propriété 2 des fonctionnelles c-conjuguées, nous pouvons conclure que:

$$a(u_0) = a^{cc}(u_0).$$

b) $a(u_0)$ est infini

. Si $a(u_0) = -\infty$, $a^{cc}(u_0) = a(u_0)$ (cfr propriété 2 des fonctionnelles c-conjuguées).

. Si $a(u_0) = +\infty$, nous pouvons démontrer que $a(u_0) = a^{cc}(u_0)$ en répétant l'argument de la démonstration du cas a) mais en considérant $\frac{1}{\varepsilon}$ au lieu de $a(u_0) - \varepsilon$ lors de l'application de la s.c.i de a en u_0 (cfr inégalité (1))

Le raisonnement est alors tout à fait similaire.

Dans une deuxième étape, nous allons établir la seconde partie de la thèse c'est à dire :

Si c appartenant à $\mathbb{R}^{U\times V}$ est pointue en un élément u_0 de U et telle que $c(\cdot,v)$ est s.c.i en u_0 pour tout v de V alors toute fonctionnelle a de $\bar{\mathbb{R}}^U$ c-tempérée satisfait :

$$\bar{a}(u_0) = a^{cc}(u_0)$$

Soit a une fonctionnelle arbitraire de $\bar{\mathbb{R}}^U$ qui est c-tempérée.

Par sa définition, $\bar{a}$ est s.c.i en tout point de U.

Il nous suffit donc de vérifier que $\bar{a}$ est c-tempérée et par la première étape de la démonstration nous aurons immédiatement la thèse.

Voyons que $\bar{a}$ est c-tempérée.

Par hypothèse, a est c-tempérée, donc il existe un élément v de V et un réel η tels que :

$$c(u,v) + \eta \leq a(u) \qquad \forall u \in U$$

Soit u' un point arbitraire de U, nous avons :

$$\liminf_{u\to u'} [c(u,v)+\eta] \leq \liminf_{u\to u'} [a(u)]$$

En utilisant la s.c.i de $c(\cdot,v)$ et la définition de $\bar{a}$ nous obtenons :

$$c(u',v) + \eta \leq \bar{a}(u')$$

Cette inégalité est vérifiée pour un élément u' arbitraire dans U donc $\bar{a}$ est c-tempérée.

Démonstration du théorème II.2

. La première partie de la thèse est immédiate, soit.

Si c est une fonctionnelle de $\bar{\mathbb{R}}^{U\times V}$ pointue en chaque élément de U alors toute fonctionnelle a de $\bar{\mathbb{R}}^U$, s.c.i et c-tempérée appartient à $\Gamma^c(U)$

En effet, en chaque point de U nous vérifions les hypothèses du théorème I.1 donc toute fonctionnelle a de $\bar{\mathbb{R}}^U$ s.c.i et c-tempérée coïncidera avec a^{cc}. Par conséquent, suite à la définition de a^{cc}, a appartiendra à $\Gamma^c(U)$.

Nous devons encore démontrer que :

Si c est une fonctionnelle de $\bar{\mathbb{R}}^{U\times V}$ pointue en chaque élément de U et telle que $c(.,v)$ est s.c.i sur U alors, pour toute fonctionnelle a de $\bar{\mathbb{R}}^U$ ne prenant jamais la valeur $(-\infty)$ nous avons que :

a appartient à $\Gamma^c(U)$ si et seulement si a est s.c.i et c-tempérée.

condition nécessaire :

Soit une fonctionnelle a arbitraire de $\bar{\mathbb{R}}^U$ appartenant à $\Gamma^c(U)$ telle que : $a(u) \neq -\infty \quad \forall u \in U$

Par la propriété 3 des fonctionnelles c-conjuguées, cela entraîne : $\quad a(u) = a^{cc}(u) \quad \forall u \in U$

donc, $\quad a^{cc}(u) \neq -\infty \quad \forall u \in U$

Et par conséquent, a est c-tempérée (cfr proposition II.1)

Il nous reste à prouver que a est s.c.i sur U.

Par hypothèse, a appartient à $\Gamma^c(U)$ donc, il existe un ensemble d'indices I, une famille d'éléments de V $\{v_i\}_{i\in I}$ et une famille de nombres réels $\{\eta_i\}_{i\in I}$ tels que.

$$a(u) = \sup_{i\in I} \left(c(u, v_i) + \eta_i \right) \qquad \forall u \in U$$

Considérons u un élément arbitraire de U.

Deux cas sont possibles

1) $a(u)$ est fini

Dans ce cas, considérons un réel $\varepsilon > 0$ quelconque

Nous allons démontrer qu'il existe un voisinage V de u tel que:

$$a(u) - \varepsilon \leq a(u') \qquad \forall u' \in V$$

Par définition de $a(u)$, nous avons:

$$a(u) - \varepsilon < \sup_{i\in I} \left(c(u, v_i) + \eta_i \right)$$

Donc, il existe un indice j dans I tel que:

$$a(u) - \varepsilon < c(u, v_j) + \eta_j$$

Par la s.c.i de $c(\cdot, v_j)$ il existe un voisinage de u,

appelons-le V tel que:

$$a(u) - \varepsilon \leq c(u', v_j) + \eta_j \qquad \forall u' \in V$$

Par conséquent :

$$a(u) - \varepsilon \leq \sup_{i\in I} \left(c(u', v_i) + \eta_i \right) = a(u') \qquad \forall u' \in V$$

a est donc s.c.i en u.

2) $a(u) = -\infty$, a est immédiatement s.c.i en u car $a(u) \leq \liminf_{u' \to u} a(u')$

3) $a(u) = +\infty$

Dans ce cas, on répète l'argument de la démonstration du cas 1) en prenant $\frac{1}{\varepsilon}$ à la place de $a(u) - \varepsilon$ et nous pouvons toujours conclure que a est s.c.i en u.

<u>condition suffisante</u>

Elle a déjà été établie dans la première partie de la démonstration.

Démonstration du théorème II.3

Soit c une fonctionnelle de $\overline{\mathbb{R}}^{U \times V}$ telle que, pour tout élément v de V, $c(\cdot, v)$ est s.c.s en u_0 appartenant à U et pour toute fonctionnelle a de $\overline{\mathbb{R}}^{U}$ s.c.i en u_0 et c-tempérée nous avons : $a(u_0) = a^{cc}(u_0)$.

Nous devons démontrer que c est pointue en u_0.

Considérons v un élément arbitraire de V, N un voisinage quelconque de u_0 et deux réels η, ε arbitraires, ε étant strictement positif.

Prouver la thèse revient à établir l'assertion suivante :

$\exists\, \hat{v} \in V$, N' voisinage de u_0, $N' \subset N$ tels que :

$$\begin{cases} c(u, \hat{v}) - c(u_0, \hat{v}) \leq \varepsilon & \forall u \in N' \quad (1) \\ c(u, \hat{v}) - c(u_0, \hat{v}) \leq c(u, v) + \eta & \forall u \notin N' \quad (2) \end{cases}$$

Nous allons envisager deux cas :

1) $c(u_0, v) + \eta \geq 0$

Dans ce cas, choisissons $\hat{v} = v$

Nous allons nous servir de la s.c.s de $c(\cdot, v)$ en u_0 pour déterminer N'.

En effet, $-c$ est s.c.i donc, il existe un voisinage $\hat{N}$ de u_0 tel que : $\quad -c(u, v) \geqslant -c(u_0, v) - \varepsilon \quad \forall u \in \hat{N}$

Posons $N' = \hat{N} \cap N$, N' est non vide.

Pour ce choix de $\hat{v}$ et de N', les inégalités (1) et (2) sont vérifiées.

2) $c(u_0, v) + \eta < 0$

Comme $c(\cdot, v)$ est s.c.s en u_0, il existe un voisinage de u_0

soit $\hat{N}$ tel que : $\quad c(u,v) + \eta < 0 \qquad \forall u \in \hat{N} \quad (3)$

Posons de nouveau $N' = \hat{N} \cap N$ et considérons une fonction f de $\overline{\mathbb{R}}^U$ définie par,

$$f(u) = \begin{cases} -\eta + \varepsilon & \text{si } u \in N' \\ c(u,v) & \text{sinon} \end{cases}$$

f possède les propriétés suivantes :

- $c(u,v) \leq f(u) \quad \forall u \in U$, elle est donc c-tempérée
- f est s.c.i en u_0 car elle est constante dans un voisinage de u_0.

Par conséquent, $f(u_0) = f^{cc}(u_0)$

Donc, $$f^{cc}(u_0) = \sup_{\substack{v \in V \\ \eta \in \mathbb{R} \\ c(u,v)+\eta \leq f(u)\ \forall u \in U}} \{ c(u_0,v) + \eta \} = -\eta + \varepsilon$$

Il s'ensuit, par définition du supremum, qu'il existe un élément v' de V et un réel η' tels que,

$$c(u,v') + \eta' \leq f(u) \qquad \forall u \in U \qquad (4)$$
$$c(u_0,v') + \eta' = \gamma > -\eta \quad , \quad \gamma \in \mathbb{R} \qquad (5)$$

Posons $\hat{v} = v'$, $\hat{\eta} = \eta' - (\gamma + \eta) = -c(u_0,\hat{v}) - \eta$

Par (4) et (5) nous obtenons :

$$c(u,\hat{v}) + \hat{\eta} \leq f(u) \qquad \forall u \in U \qquad (6)$$

En nous référant aux définitions de f et de $\hat{\eta}$, l'expression (6) peut encore s'écrire

$$c(u,\hat{v}) - c(u_0,\hat{v}) - \eta \leq -\eta + \varepsilon \qquad \forall u \in N'$$
$$c(u,\hat{v}) - c(u_0,\hat{v}) - \eta \leq c(u,v) + \eta \qquad \forall u \notin N'$$

Dans tous les cas nous obtenons la thèse et par conséquent, c est pointue en u_0.

Démonstration du théorème II.5

Condition nécessaire

Supposons que (P) est inf-stable c'est à dire $h(\sigma)$ est fini et h est s.c.i en σ.

Par le théorème II.4 nous avons immédiatement $\beta = -\alpha$

Condition suffisante

Supposons $-\beta = \alpha$, nous devons vérifier que h est s.c.i en σ.

Remarquons que $h(\sigma)$ est bien fini car $h(\sigma) = \alpha = -\beta$ fini par hypothèse.

Par la formule (II.1) (page 65) nous avons : $h^{cc}(\sigma) = h(\sigma)$

En outre, les hypothèses de la deuxième partie du théorème II.1 sont vérifiées donc $h(\sigma) = \bar{h}(\sigma)$.

Or, par définition $\bar{h}$ est s.c.i sur U et minore h.

Par conséquent, si nous considérons un réel arbitraire $\varepsilon > 0$, il existe un voisinage V de σ tel que :

$$\bar{h}(\sigma) - \varepsilon \leq \bar{h}(u) \qquad \forall u \in V$$

donc, $\quad h(\sigma) - \varepsilon \leq \bar{h}(u) \leq h(u) \qquad \forall u \in V$

Nous en déduisons que h est s.c.i en σ.

ANNEXE III

Démonstration du théorème II.6

Soit $(\bar{x}, \bar{v})$ appartenant à $X \times V$, un point de selle du lagrangien, nous devons démontrer que $\bar{v}$ est une solution du problème dual c'est à dire que $g(\bar{v}) = \beta$

Par définition de g et du lagrangien nous avons:

$$g(\bar{v}) = \sup_{x \in X} l(x, \bar{v})$$

Par conséquent, puisque $(\bar{x}, \bar{v})$ est un point de selle de l

$$g(\bar{v}) = l(\bar{x}, \bar{v}) = \inf_{v \in V} l(\bar{x}, v) \qquad (1)$$

donc, $g(\bar{v}) \leq \inf\limits_{v \in V} \sup\limits_{x \in X} l(x, v) = \beta$

or, $\beta = \inf\limits_{v \in V} g(v)$

Il s'ensuit que $g(\bar{v}) = \beta \qquad (2)$

Supposons, en outre, que $-\alpha = \beta < +\infty$, et c pointue en σ. Dans ce cas, si $\phi(\bar{x}, .)$ est s.c.i en σ alors $\bar{x}$ est une solution du problème primal c'est à dire $f(\bar{x}) = \alpha$.

Pour vérifier cette assertion nous allons prouver que $\phi(\bar{x}, .)$ est c-tempérée et en déduire la thèse en appliquant le théorème II.1.

Voyons que $\phi(\bar{x},.)$ est c-tempérée. Pour cela, nous allons calculer $\phi^{cc}(\bar{x},\sigma)$.

$$\phi^{cc}(\bar{x},\sigma) \overset{\text{déf}}{=} \sup_{v\in V} [c(\sigma,v) - \phi^c(\bar{x},v)]$$

Suite à la normalisation de c en σ et à la définition de $\phi^c(\bar{x},v)$ nous avons :

$$\phi^{cc}(\bar{x},\sigma) = -\inf_{v\in V} \sup_{u\in U} [c(u,v) - \phi(\bar{x},u)]$$

$$= -\inf_{v\in V} l(\bar{x},v)$$

Par (1) et (2) nous obtenons donc :

$$\phi^{cc}(\bar{x},\sigma) = -\beta \qquad (3)$$

Par conséquent, $\phi^{cc}(\bar{x},\sigma)$ est fini ce qui implique, par la proposition II.1 que $\phi(\bar{x},.)$ est c-tempérée.

Toutes les hypothèses du théorème II.1 sont remplies pour $\phi(\bar{x},.)$ donc, $\phi(\bar{x},\sigma) = \phi^{cc}(\bar{x},\sigma) = -\beta = \alpha$

comme $\phi(\bar{x},\sigma) \overset{\text{déf}}{=} f(\bar{x})$ nous avons $f(\bar{x}) = \alpha$.

Il nous reste à prouver l'équivalence entre la semi-continuité inférieure de $\phi(\bar{x},.)$ et le fait que $f(\bar{x}) = \alpha$ si nous avons l'hypothèse supplémentaire "$c(.,v)$ est s.c.i en σ pour tout élément v de V".

<u>condition suffisante</u>

Elle est immédiate par l'étape précédente.

<u>condition nécessaire</u>

Soit $f(\bar{x}) = \alpha$, $\phi(\bar{x},.)$ vérifie maintenant toutes les hypothèses de la seconde partie du théorème II.1 en σ.

Il s'ensuit :

$$\phi(\bar{x},\sigma) = \phi^{cc}(\bar{x},\sigma)$$

donc, par (3) et grâce à l'hypothèse $\beta = -\alpha$, nous déduisons :

$$\phi(\bar{x},\sigma) = \alpha = f(\bar{x}) = \phi(\bar{x},\sigma)$$

Par un raisonnement semblable à celui réalisé dans la démonstration du théorème II.5 (annexe II) nous concluons que $\phi(\bar{x},\cdot)$ est s.c.i en σ

Démonstration du théorème II.7

Soit $c\text{-}\partial h(\sigma) \neq \phi$.

Nous devons prouver que si $\bar{x}$ est une solution de (P) et $\bar{v}$ est une solution de (Q), alors le couple $(\bar{x},\bar{v})$ est un point de selle du lagrangien.

Et, en particulier, pour $\bar{v}$ appartenant à $c\text{-}\partial h(\sigma)$, nous avons :

$$\sup_{x\in X} l(x,\bar{v}) = -\alpha$$

Vérifions cette assertion :

Suite à la définition de l et à l'hypothèse sur $(\bar{x},\bar{v})$ nous obtenons les inégalités suivantes :

$$l(x,\bar{v}) \leq \sup_{x\in X} l(x,\bar{v}) = g(\bar{v}) = \beta \qquad \forall x\in X \qquad (1)$$

$$-\alpha = -f(\bar{x}) \leq \inf_{v\in V} l(\bar{x},v) \qquad (2)$$

Comme $c\text{-}\partial h(\sigma) \neq \phi$ cela implique : (cfr proposition II.4)

$$\beta = -\alpha \qquad (3)$$

En réunissant les résultats (1), (2) et (3) nous obtenons :

$$l(x,\bar{v}) \leq \inf_{v\in V} l(\bar{x},v) \leq l(\bar{x},\bar{v}) \leq \sup_{x\in X} l(x,\bar{v}) \leq \inf_{v\in V} l(\bar{x},v) \leq l(\bar{x},v), \quad \forall x\in X, \ \forall v\in V$$

Par conséquent, $(\bar{x}, \bar{v})$ est un point de selle du lagrangien.

En particulier, si $\bar{v}$ appartient à c-$\partial h(\sigma)$, en utilisant la proposition II.4 et la propriété 3 des c-sous-différentiels nous déduisons que $\bar{v}$ appartient à c-$\partial h^{cc}(\sigma)$.
Par conséquent, $\bar{v}$ appartient à B (cf. proposition II.3)
donc $g(\bar{v}) = \sup\limits_{x \in X} l(x, \bar{v}) = \beta = -\alpha$

ANNEXE IV

Démonstration de l'inf-stabilité des programmes mathématiques (théorème III.1)

Par hypothèse, nous savons déjà que $h(o) = \alpha$ est fini, il reste à démontrer que h est s.c.i en o.

Pour cela, nous allons appliquer le lemme suivant :

Lemme

Soit (f_α), α appartenant à A, une famille de fonctionnelles.

A étant un compact quelconque.

Posons $f = \underset{\alpha \in A}{\mathrm{Inf}}\, f_\alpha$

Supposons qu'il existe un ouvert $\mathcal{U}$ de X tel que l'application

$$(\alpha, x) \in A \times \mathcal{U} \longrightarrow f_\alpha(x) \quad \text{soit s.c.i sur } A \times \mathcal{U}$$

Alors, en tout point x_0 de $\mathcal{U}$ la fonctionnelle f est s.c.i

Démonstration du lemme

Soit x_0, un élément arbitraire de $\mathcal{U}$.

Si $f(x_0) = -\infty$ f est automatiquement s.c.i en x_0

Si $f(x_0) \neq -\infty$, considérons ε, un réel strictement positif arbitraire.

Nous allons démontrer qu'il existe un voisinage $\mathcal{V}$ de x_0 tel que :

$$f(x_0) - \varepsilon \leq f(x) \quad \forall x \in \mathcal{V} \text{ si } f(x_0) \text{ est fini}$$

ou

$$\frac{1}{\varepsilon} \leq f(x) \quad \forall x \in \mathcal{V} \text{ si } f(x_0) = +\infty$$

Nous ferons la démonstration pour $f(x_0)$ fini, l'autre cas est similaire

Par définition de f, $f(x_0) - \varepsilon \leq f_\alpha(x_0) \quad \forall \alpha \in A$

Comme la fonctionnelle $(\alpha, x) \in A \times \mathcal{U} \to f_\alpha(x)$ est s.c.i sur $A \times \mathcal{U}$,

$\forall \alpha \in A$, $\exists\, \mathcal{O}_\alpha$ voisinage ouvert de α, $\exists\, \mathcal{V}_\alpha$ voisinage de x_0 tels que :

$\forall \alpha' \in O_\alpha$, $\forall x \in V_\alpha$ $\qquad f(x_0) - \varepsilon \leq f_{\alpha'}(x)$

Or A est compact, et A est inclus dans $\bigcup_{\alpha \in A} O_\alpha$, par conséquent,

$\exists \alpha_1 \ldots \alpha_m \in A$ tels que : $\left(\bigcup_{i=1,m} O_{\alpha_i} \right) \supseteq A$ $\qquad (m \in \mathbb{N})$

Posons $V = \bigcap_{i=1 \ldots m} V_{\alpha_i}$ nous obtenons

$\forall x \in V$, $\forall \alpha' \in A$ $\quad f(x_0) - \varepsilon \leq f_{\alpha'}(x)$

donc,

$\forall x \in V$ $\qquad f(x_0) - \varepsilon \leq \operatorname{Inf}_{\alpha \in A} f_\alpha(x) = f(x)$

Il s'ensuit que f est s.c.i en tout x_0 de $\mathcal{U}$

<u>Remarque</u> : Si nous satisfaisons uniquement l'hypothèse "$f(u) = \operatorname{Inf}_{\alpha \in A} f_\alpha(u)$" pour tout u appartenant à un voisinage de x_0 nous pourrons conclure que f est s.c.i en x_0 si les autres hypothèses sont satisfaites.
Ceci est évident à l'étude de la démonstration précédente.

Dans notre cas, nous allons démontrer qu'il existe un voisinage V de o tel que :

$$h(u) = \inf_{x \in K} \phi(x,u) \qquad \forall u \in V$$

$$\text{où } K = \{ x \in X \mid G(x) \leq -\bar{u} \text{ et } f_0(x) \leq \bar{\alpha} \} \quad \text{compact}$$

et nous démontrerons ensuite la semi-continuité inférieure de la fonctionnelle ϕ sur $K \times V$.

Nous supposons que l'intérieur de C est non vide, la démonstration pour le cas où l'intérieur de C est vide découlera de façon immédiate de celle ci-dessous, il suffira de supprimer certains passages qui ne seraient plus nécessaires dans ce cas.

1) Par hypothèse, $\bar{u}$ appartient à l'intérieur de C donc, il existe un ouvert $\mathcal{O}$ tel que :

$\bar{u} \in \mathcal{O}$ et $\mathcal{O} \subseteq C$.

Considérons $W = \mathcal{O} - \bar{u}$, c'est un voisinage de 0 car 0 appartient à W et en outre, c'est la translation d'un ouvert dans l'espace vectoriel topologique U.

Posons $V = W \cap \{u \in U \mid \exists x \in X$ vérifiant $G(x) \leq u$ et $f_0(x) \leq \bar{z}\}$

Nous allons prouver que:

$$h(u) = \inf_{x \in K} \phi(x, u) \qquad \forall u \in V$$

Considérons un élément arbitraire u de V, par définition de W il existe un élément u_1 de $\mathcal{O}$ tel que : $u = u_1 - \bar{u}$

Par conséquent $u + \bar{u}$ appartient à $\mathcal{O}$ et donc à C.

Pour tout x de X vérifiant $G(x) \leq u$ ($\Leftrightarrow G(x) - u \in C$) nous avons, puisque C est un cône convexe que:

$$G(x) - u + u + \bar{u} \in C$$

donc $G(x) \leq -\bar{u}$.

Or, $h(u) = \inf\limits_{x \in X} \phi(x, u) = \inf\limits_{\{x \mid G(x) \leq u\}} \phi(x, u) \qquad \forall u \in U$

(car en dehors de $\{x \mid G(x) \leq u\}$ $\phi(x, u)$ vaut $+\infty$)

Il s'ensuit:

$$h(u) = \inf_{\{x \mid G(x) \leq -\bar{u}\}} \phi(x, u) \qquad \forall u \in V.$$

Nous savons également que pour un élément u de V il existe $x \in X$ vérifiant $G(x) \leq u$ et $f_0(x) \leq \bar{\alpha}$ ce qui revient à $\phi(x,u) \leq \bar{\alpha}$.
Par conséquent, nous pouvons encore limiter l'ensemble sur lequel nous prenons l'infimum soit:

$$h(u) = \inf_{\{x \in X \mid G(x) \leq -\bar{u} \text{ et } f_0(x) \leq \bar{\alpha}\}} \phi(x,u) \quad \forall u \in V$$

$$h(u) = \inf_{x \in K} \phi(x,u) \qquad \forall u \in V$$

2) Il nous reste à prouver que la fonctionnelle
$(x,u) \in K \times V \longrightarrow \phi(x,u)$ est s.c.i sur $K \times V$.
Ceci revient à démontrer que si on considère un élément (x,u) arbitraire dans $K \times V$ alors

$$\phi(x,u) \leq \liminf_{(x',u') \to (x,u)} \phi(x',u')$$

Il nous suffit donc de vérifier que pour tout système dirigé $\{(x_i, u_i)\}_{i \in I} \subseteq K \times V$ convergeant vers (x,u) si la limite de $\{\phi(x_i,u_i)\}_{i \in I}$ existe dans $\bar{\mathbb{R}}$ alors $\phi(x,u) \leq \lim \phi(x_i,u_i)$ (1)

Considérons donc un système dirigé arbitraire $\{(x_i,u_i)\}_{i \in I}$ tel que $\{\phi(x_i,u_i)\}_{i \in I}$ admette une limite

Trois cas sont possibles:

① $\exists j \in I$ tel que $\forall i > j \quad G(x_i) \leq u_i$
Puisque G est fermé pour l'ordre nous obtenons $G(x) \leq u$
ce qui implique : $\phi(x,u) = f_0(x)$
En outre, $\lim \phi(x_i,u_i) = \lim f_0(x_i) = \liminf f_0(x_i)$
Or, f_0 est s.c.i en x donc $f_0(x) = \liminf f_0(x_i)$
Par conséquent, $\phi(x,u) = \lim \phi(x_i,u_i)$

② $\exists j \in I$ tel que $\forall i > j$ $G(x_i) \nleq u_i$
c'est-à-dire $\phi(x_i, u_i) = +\infty$ $\forall i > j$
nous aurons donc $\lim \phi(x_i, u_i) = +\infty$ et l'inégalité (1)
sera bien vérifiée.

③ $\forall j \in I$ tels que $G(x_j) \leq u_j$ $\exists i > j$ tel que $G(x_i) \nleq u_i$ et
$\forall j \in I$ tels que $G(x_j) \nleq u_j$ $\exists i > j$ tel que $G(x_i) \leq u_i$
Il s'ensuit :
$\forall i \in I$ $\exists j > i$ tel que $\phi(x_i, u_i) = +\infty$
Puisque nous supposons que la limite de $\{\phi(x_i, u_i)\}_{i \in I}$
existe, elle ne peut être que $+\infty$ et par conséquent,
l'inégalité (1) est à nouveau bien vérifiée.

Toutes les hypothèses du lemme sont satisfaites dans un voisinage de σ. Nous pouvons l'appliquer et conclure que la fonction de perturbation h est s.c.i en σ.

ANNEXE V

La méthode de pénalisation est une procédure pour approcher des problèmes d'optimisation contraints par des problèmes non contraints.
Cette approximation est accomplie en additionnant à la fonction objective à minimiser un terme prenant une grande valeur à l'extérieur du domaine défini par les contraintes.
Associé à cette méthode, nous avons un paramètre M qui détermine la sévérité de la pénalisation et, par conséquent, le degré d'approximation du problème contraint par le problème non contraint. Quand nous faisons tendre M vers $(+\infty)$, l'approximation devient de plus en plus précise.

Deux questions fondamentales sont associées à cette méthode :

1º De quelle manière le problème non contraint approche-t-il le problème contraint ?
Pour répondre à cette question, il faudrait analyser le comportement de la solution du problème non contraint quand M tend vers $(+\infty)$ et voir si elle tend vers la solution du problème contraint.

2º Comment résoudre un problème non contraint quand sa fonction objective contient un paramètre de pénalisation ?
Il s'avère, en fait, que la croissance de M nécessaire pour avoir un bon problème approximant implique un conditionnement de plus en plus mauvais du hessien correspondant à la fonction objective du problème non contraint. Par conséquent, la vitesse de convergence de la plupart des algorithmes appliqués pour

résoudre les problèmes non contraints est de beaucoup ralentie. (Pour plus de détails voir [5] page 278)

Par lagrangien généralisé de la méthode de pénalisation extérieure nous désignons, en fait, un lagrangien augmenté de la méthode des multiplicateurs de Hestenes et Powell. Ces derniers ont synthétisé en une seule méthode deux techniques de résolution de problèmes contraints en essayant d'éliminer les inconvénients de chacune tout en conservant les avantages.

Donnons un résumé très bref de ces trois méthodes.

Pour plus de facilité, nous envisageons le cas particulier du problème contraint suivant :

$$\left.\begin{array}{l} \text{minimiser } f(x) \\ \text{sous contraintes : } x \in X,\ h_1(x) = h_2(x) = \dots h_m(x) = 0 \end{array}\right\} \quad (1)$$

où $f, h_i : \mathbb{R}^n \to \mathbb{R}$ et X est un sous-ensemble de $\mathbb{R}^n$.

1) <u>Méthodes de pénalisation.</u>

L'idée est de procéder à des minimisations séquentielles de la forme :

$$\inf_{x \in X} \left[f(x) + c_k \cdot \sum_{i=1}^{m} \phi[h_i(x)] \right] \qquad (2)$$

où $\{c_k\}_{k \in \mathbb{N}}$ est une suite de nombres réels telle que :

- $c_k \leq c_{k+1} \quad \forall k \in \mathbb{N}$
- $\{c_k\}_{k \in \mathbb{N}}$ tend vers $(+\infty)$ pour k croissant.

et ϕ est une fonction de $\mathbb{R}$ vers $[0, +\infty]$ vérifiant :

$$\left.\begin{array}{ll} \cdot\ \phi(t) \geq 0 & \forall t \in \mathbb{R} \\ \cdot\ \phi(t) = 0 & \text{si } t = 0 \end{array}\right\} \quad (3)$$

exemple : $\phi(t) = \frac{1}{2} t^2$ satisfait les conditions (3).

Cette suite de minimisations engendre une suite de points x_k de X (x_k étant le minimum de (2)) et une suite $\tilde{y}_k$ de réels où $\tilde{y}_k = (c_k\, \phi'[h_1(x_k)], \ldots\ c_k\, \phi'[h_m(x_k)])$ (si nous supposons l'existence de $\phi' = \frac{d\phi}{dt}$).

Considérons ensuite la limite suivante :

$$\lim_{c_k \to \infty} \inf_{x \in X} \{ f(x) + c_k \sum_{i=1}^{m} \phi[h_i(x)] \} \qquad (4)$$

Or, par (3) la valeur optimale de (1) peut s'écrire :

$$\inf_{x \in X} \lim_{c_k \to \infty} \{ f(x) + c_k \sum_{i=1}^{m} \phi[h_i(x)] \} \qquad (5)$$

Le <u>succès de la méthode</u> est donc basé sur l'égalité des expressions (4) et (5).

(hypothèses nécessaires voir [5], page 278_280)

En outre, la continuité de f, h_i et ϕ près de la solution sont des hypothèses garantissant l'existence d'une solution x_k du problème (2) (cfr. [5]).

De plus, si ϕ est dérivable la suite $\{\tilde{y}_k\}$ converge sous certaines hypothèses vers le multiplicateur de Lagrange du problème ([5], page 285).

<u>Côtés négatifs de la méthode</u> : Nous avons une convergence lente et des instabilités numériques associées au mauvais conditionnement du problème (2), celui-ci étant induit par les grandes valeurs du paramètre de pénalisation c_k.

2) Méthodes basées sur des minimisations séquentielles du lagrangien.

Nous définissons le lagrangien du problème (1) par:

$$L(x,y) = f(x) + \sum_{i=1}^{m} y^i h_i(x) \qquad x \in \mathbb{R}^n,\ y = (y^1, \cdots y^m) \in \mathbb{R}^m \quad (6)$$

L'idée est de minimiser $L(x, y_k)$ sur $x \in X$ pour une suite de vecteurs de multiplicateurs $\{y_k\}_{k \in N}$.
Cette suite est engendrée par l'itération suivante:

$$y^i_{k+1} = y^i_k + \alpha_k\, h_i(x_k) \qquad i = 1, \dots m \qquad (7)$$

où x_k est le minimum de $L(x, y_k)$ sur $x \in X$ et α_k est un scalaire (paramètre donnant la taille du pas)
Remarquons que (7) peut-être considérée comme une itération de plus grande descente associée à la recherche de la solution du dual.
C'est pourquoi cette méthode est appelée méthode pimal-dual.

Désavantages de la méthode.

① Le problème (1) doit posséder une structure localement convexe afin que le dual soit bien défini et que l'itération (7) ait un sens. ([5], p 312)

② Il est souvent nécessaire de faire de nombreuses minimisations du lagrangien avant d'atteindre la solution car l'itération de descente (7) ne converge seulement que modérément vite.

Par conséquent, cette méthode est appliquée uniquement dans le cas particulier où les minimisations successives du lagrangien aboutissent vite suite à certaines structures spéciales du problème initial. ([5])

3) Méthode des multiplicateurs

Cette méthode reprend des idées de chacune des précédentes. Ici, nous considérons un lagrangien augmenté

$$L(x,y,c) = \underbrace{f(x) + \sum_{i=1}^{m} y^i h_i(x)}_{\text{partie habituelle du lagrangien}} + \underbrace{c \sum_{i=1}^{m} \phi[h_i(x)]}_{\text{fonction de pénalisation}} \qquad (8)$$

y correspond implicitement à des multiplicateurs et c à une constante de pénalisation, ϕ vérifiant toujours les conditions (3).

Nous procédons de nouveau à une minimisation séquentielle du lagrangien soit :

$$\inf_{x \in X} L(x, y_k, c_k) = f(x) + \sum_{i=1}^{m} y_k^i h_i(x) + c_k \sum_{i=1}^{m} \phi[h_i(x)] \qquad (9)$$

$\{c_k\}_{k \in \mathbb{N}}$ étant une suite de paramètres de pénalisation positifs et $\{y_k\}_{k \in \mathbb{N}}$ une suite de multiplicateurs engendrée par l'itération suivante :

$$y_{k+1}^i = y_k^i + c_k \, \phi'[h_i(x_k)] \qquad i = 1, \dots m \qquad (10)$$

(Nous faisons l'hypothèse que $\phi' = \dfrac{d\phi}{dt}$ existe)

où x_k est le point minimum de $L(x, y_k, c_k)$; y_0 est un point sélectionné a priori et la suite $\{c_k\}_{k \in \mathbb{N}}$ est soit présélectionnée, soit engendrée au cours du calcul suivant un certain schéma (cfr [6])

Remarque : Habituellement $\phi(t) = \frac{1}{2} t^2$.

Dans ce cas, l'expression (10) devient :

$$y_{k+1}^i = y_k^i + c_k \, h_i(x_k) \quad \text{un cas particulier de (7).}$$

Supposons que $\{c_k\}_{k\in\mathbb{N}}$ tende vers $(+\infty)$ et que $\{y_k\}_{k\in\mathbb{N}}$ soit une suite bornée.

Dans ce cas, la méthode donne la valeur optimale de (1) moyennant certaines hypothèses qui garantisent la validité de l'échange "lim" et "inf" dans l'expression suivante :

$$\lim_{c_k\to\infty} \inf_{x\in X} \{ f(x) + \sum_{i=1}^{m} y_k^i\, h_i(x) + c_k \sum_{i=1}^{m} \phi[h_i(x)] \}$$

<u>Avantages</u>

Nous pouvons maintenant avoir la convergence sans nécessairement faire croître c_k jusque $(+\infty)$, c'est à dire la convergence n'est pas simplement déduite par des valeurs toujours croissantes du paramètre de pénalisation mais aussi par l'itération (10) sur le vecteur des multiplicateurs.

Par conséquent, le mauvais conditionnement du hessien lié à la méthode de pénalisation peut en partie être évité.

En outre, l'itération (10) converge rapidement vers un vecteur multiplicateur de Lagrange du problème (1) sous des hypothèses relativement faibles, plus vite que dans la méthode primal-dual considérée plus tôt et cela sans que le problème (1) ait spécialement une structure localement convexe.

(Pour compléments sur la méthode des multiplicateurs voir [6])

ANNEXE VI

Calcul des lagrangiens des exemples 1-b et 2-b

Exemple 1.b

$U = \mathbb{R}^m \quad , V = \mathbb{R}_+^m$

$c(u,v) = -\sum_{i=1}^{m} v_i |u_i| \qquad , u = (u_1, \dots u_m) \in \mathbb{R}^m, v = (v_1 \dots v_m) \in \mathbb{R}_+^m$

$G = (g_1, \dots g_p, g_{p+1}, \dots g_m) \qquad 1 \leq p \leq m$

où $g_i : X \to \mathbb{R} \quad i = 1 \dots m$

$C = \mathbb{R}_-^p \times \{ 0_{\mathbb{R}^{m-p}} \}$

où $0_{\mathbb{R}^{m-p}}$ désigne l'origine de $\mathbb{R}^{m-p}$

Nous allons d'abord vérifier que nous avons bien un programme mathématique :

1°) C est un cône

Soit x appartenant à C, nous devons démontrer que λx appartient à C, pour tout réel λ positif.

En effet, $x(x_1, \dots x_m) \in C$ implique :

$$\left.\begin{array}{l} x_1, \dots x_p \leq 0 \\ x_{p+1} = \dots = x_m = 0 \end{array}\right\} \text{donc,} \quad \left.\begin{array}{l} \lambda x_1, \dots \lambda x_p \leq 0 \\ \lambda x_{p+1} = \dots = \lambda x_m = 0 \end{array}\right\} \quad \forall \lambda \geq 0$$

Ce qui entraine, par définition de C, $\quad \lambda x \in C \quad \forall \lambda \geq 0$

2°) C est convexe.

Comme C est un cône, il suffit de prouver, pour x, y arbitraires dans C, que $x + y$ appartient à C.

$$x, y \in C \Rightarrow \begin{cases} x_1, \dots x_p, y_1 \dots y_p \leq 0 \\ x_{p+1} = \dots = x_m = y_{p+1} = \dots = y_m = 0 \end{cases} \Rightarrow \begin{cases} x_1 + y_1, \dots x_p + y_p \leq 0 \\ x_{p+1} + y_{p+1} = \dots = x_m + y_m = 0 \end{cases}$$

donc $x + y \in C$.

Calculons maintenant $\bar{c}(G(x),v)$, $v \in V$, $x \in X$.

Par définition,

$$\bar{c}(G(x),v) = \operatorname{Sup}\{-\sum_{i=1}^{m} v_i |u_i| \mid u \in U,\ G(x) \leq u\}$$

Suite à la définition de l'ordre sur U, nous avons :

$$\bar{c}(G(x),v) = \operatorname{Sup}\{-\sum_{i=1}^{m} v_i |u_i| \mid u \in U,\ G(x) - u \in C\}$$

Par conséquent, grâce à l'expression de C,

$$\begin{aligned}\bar{c}(G(x),v) &= \operatorname{Sup}\{-\sum_{i=1}^{m} v_i |u_i| \mid u \in U;\ g_i(x) - u_i \leq 0 \quad 1 \leq i \leq p, \\ &\qquad \text{et } g_j(x) - u_j = 0 \quad p+1 \leq j \leq m\} \\ &= \operatorname{Sup}\{-\sum_{i=1}^{p} v_i |u_i| \mid u \in U;\ g_i(x) - u_i \leq 0 \quad 1 \leq i \leq p\} \\ &\quad - \sum_{i=1}^{p+1} v_i |g_i(x)|.\end{aligned}$$

$$\bar{c}(G(x),v) = -\sum_{i=1}^{p} v_i \max(g_i(x),0) - \sum_{i=1}^{p+1} v_i |g_i(x)|$$

En effet, comme l'on a une somme de termes négatifs, le suprémum sera atteint en donnant la valeur minimale possible à $|u_i|$ $i = 1, .. p$.

. Soit 0, mais cela n'est possible que si $g_i(x) \leq 0$ $i = 1, .. p$ (cfr condition sur u)

. Si $g_i(x) > 0$ $i = 1, .. p$; le suprémum sera atteint pour $u_i = g_i(x)$ $i = 1, .. p$.

<u>Exemple 2.b</u>

$U = \mathbb{R}^m$, $V = \mathbb{R}_+ \times \mathbb{R}^m$

$c(u,v) = -v_0 |u|^2 - \langle w, u \rangle$, $v = (v_0, w) \in \mathbb{R}_+ \times \mathbb{R}^m$

$u = (u_1, .. u_m) \in \mathbb{R}^m$

$G = (g_1, \dots g_p, g_{p+1}, \dots g_m)$ $\quad 1 \leq p \leq m$ où $g_i : X \to \mathbb{R}$ $\quad i = 1, .. m$

$C = \mathbb{R}^p_- \times \{0_{\mathbb{R}^{m-p}}\}$

De nouveau, nous avons bien que C est un cône convexe.

Calculons $\bar{c}(G(x), v)$, $x \in X$, $v \in V$.

$$\bar{c}(G(x), v) \overset{\text{déf}}{=} \operatorname{Sup}\{-v_0 |u|^2 - \langle w, u\rangle \mid u \in U,\ G(x) \leq u\}$$

$$= \operatorname{Sup}\{-v_0|u|^2 - \langle w, u\rangle \mid u \in U\ ;\ g_i(x) - u_i \leq 0,\ 1 \leq i \leq p$$
$$\text{et } g_j(x) = u_j\ \ p+1 \leq j \leq m\}$$

Par définition de la norme et du produit scalaire dans $\mathbb{R}^m$, nous déduisons :

$$\bar{c}(G(x), v) = \operatorname{Sup}\{-v_0 \sum_{i=1}^{p} u_i^2 - \sum_{i=1}^{p} w_i . u_i \mid u \in U, g_i(x) - u_i \leq 0,\ 1 \leq i \leq p\}$$

$$- \sum_{i=p+1}^{m} [v_0 (g_i(x))^2 + w_i\, g_i(x)].$$

Nous allons chercher pour $i = 1 .. p$

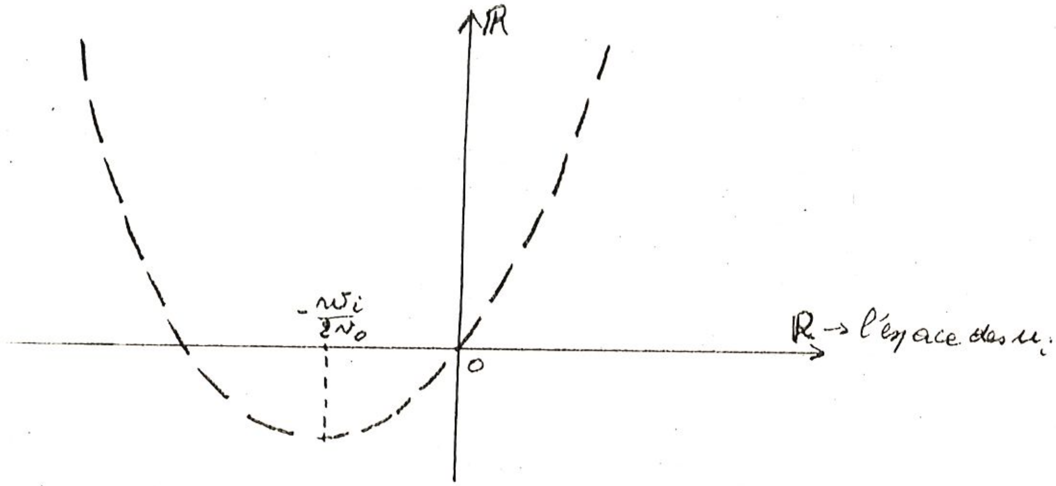
$$\begin{cases} \inf\limits_{u_i \in \mathbb{R}} [v_0 u_i^2 + w_i u_i] & u_i \in \mathbb{R},\ v_0 \in \mathbb{R}_+,\ w_i \in \mathbb{R} \\ \text{sous la contrainte } g_i(x) - u_i \leq 0 \end{cases}$$

Si nous parvenons à déterminer u_i ($i = 1, .. p$) solution de ce problème nous aurons trouvé la valeur du supremum.

ⓐ Si $v_0 = 0$, la fonction à minimiser est une droite passant par l'origine, de coefficient angulaire w_i.
Par conséquent,

- si $w_i > 0$, cette fonction est croissante et le minimum en tenant compte des contraintes sera atteint pour $u_i = g_i(x)$
- si $w_i < 0$, cette fonction est décroissante et le minimum sur $\mathbb{R}$ sera atteint pour $u_i = -\infty$.
- si $w_i = 0$, tout réel u_i supérieur ou égal à $g_i(x)$ convient.

b) Si $v_0 \neq 0$, $v_0 u_i^2 + w_i u_i$ (v_0, w_i fixés) est une parabole dont le sommet est atteint en $-\frac{w_i}{2v_0}$.
Ce sommet est un minimum de la fonction car $v_0 > 0$



Nous envisageons deux cas:

- si $g_i(x) \leq -\frac{w_i}{2v_0}$, dans ce cas, le minimum sera atteint en $-\frac{w_i}{2v_0}$
- si $g_i(x) > -\frac{w_i}{2v_0}$, le minimum sera atteint en $g_i(x)$

Dans tous les cas, le minimum sera atteint pour:

$$u_i = \max\left(g_i(x), -\frac{w_i}{2v_0}\right)$$

(si $v_0 = 0$, $-\frac{w_i}{2v_0}$ vaut $+\infty$ ou $-\infty$ suivant le signe de w_i, nous le posons arbitrairement égal à 1 si $w_i = 0$)

Par conséquent l'expression de $\bar{c}(G(x), v)$ devient:

$$\bar{c}(G(x),v) = -\sum_{i=1}^{p}\left[v_0 \max^2\left(g_i(x), -\frac{w_i}{2v_0}\right) + w_i \max\left(g_i(x), -\frac{w_i}{2v_0}\right)\right]$$

$$-\sum_{i=p+1}^{m}\left[v_0 (g_i(x))^2 + w_i g_i(x)\right]$$

$$x \in X, \; v = (v_0, w) \in V.$$

ANNEXE VII

Vérification de la flexibilité en σ des fonctionnelles couplantes des exemples 1.a et 2.a

Exemple 1.a

$$U = \mathbb{R}^m,\ V = \mathbb{R}_+^m \quad ; \ \sigma = (0,0,\dots 0) \in \mathbb{R}^m$$

$$c(u,v) = -\sum_{i=1}^{m} v_i |u_i| ,\quad u = (u_1, \dots u_m) \in \mathbb{R}^m$$

$$v = (v_1, \dots v_m) \in \mathbb{R}_+^m .$$

Soient v', v'' arbitraires dans V et un réel η quelconque.
Soit N un voisinage de σ.
Nous devons démontrer qu'il existe un élément $\bar{v}$ de V et un voisinage N' inclus dans N tels que :

$$\begin{cases} c(u,\bar{v}) \leq c(u,v') + \eta & \forall u \notin N' \quad (1) \\ c(u,\bar{v}) \leq c(u,v'') & \forall u \in N' \quad (2) \end{cases}$$

Nous savons déjà que c est de type aiguille et est normalisée en σ.
Donc, il existe v appartenant à V et un voisinage N' de σ inclus dans N tels que :

$$\begin{cases} -\sum_{i=1}^{m} v_i |u_i| \leq -\sum_{i=1}^{m} v'_i |u_i| + \eta & \forall u \notin N' \\ -\sum_{i=1}^{m} v_i |u_i| \leq 0 & \forall u \in N' \end{cases}$$

Considérons un élément $\bar{v}$ de V tel que : $\bar{v}_i = \max(v_i, v''_i)$ $i = 1, \dots m$

Par conséquent,

$$\begin{cases} -\sum_{i=1}^{m} \bar{v}_i |u_i| \leq -\sum_{i=1}^{m} v_i |u_i| \leq -\sum_{i=1}^{m} v'_i |u_i| + \eta & \forall u \notin N' \\ -\sum_{i=1}^{m} \bar{v}_i |u_i| \leq -\sum_{i=1}^{m} v''_i |u_i| & \forall u \in N' \end{cases}$$

c est donc flexible en σ car pour $\bar{v}$ et N' les inégalités (1) et (2) sont vérifiées.

Exemple 2.a :

. $U = H$ un espace de Hilbert, $V = \mathbb{R}_+ \times H$, σ est le neutre de H pour l'addition et $\langle \cdot,\cdot \rangle$ est le produit scalaire dans H.

. $c(u,v) = -v_0|u|^2 - \langle w,u \rangle$, $v = (v_0, w) \in \mathbb{R}_+ \times H$, $u \in H$.

. Nous supposons que H est de dimension finie.

Soient v', v'' quelconques dans V, un réel η arbitraire et un voisinage N de σ.

Nous devons démontrer qu'il existe un élément $\bar{v}$ de V et un voisinage N' de σ, inclus dans N tels que :

$$\left| \begin{array}{lll} -\bar{v}_0|u|^2 - \langle \bar{w}, u \rangle \leq -v_0'|u|^2 - \langle w', u \rangle + \eta & \forall u \notin N' & (1) \\ -\bar{v}_0|u|^2 - \langle \bar{w}, u \rangle \leq -v_0''|u|^2 - \langle w'', u \rangle & \forall u \in N' & (2) \end{array} \right.$$

Comme c est de type aiguille en σ, il existe $v = (v_0, w) \in V$ et un voisinage N' de σ, $N' \subseteq N$ tels que :

$$\left| \begin{array}{ll} -v_0|u|^2 - \langle w, u \rangle \leq -v_0'|u|^2 - \langle w', u \rangle + \eta & \forall u \notin N' \\ -v_0|u|^2 - \langle w, u \rangle \leq 0 & \end{array} \right.$$

Choisissons $\bar{v} = (\bar{v}_0, w'')$ où $\bar{v}_0 \geq v_0''$, nous avons alors immédiatement (2).

Nous allons fixer $\bar{v}_0$ pour qu'il vérifie l'inégalité suivante :

$$-\bar{v}_0|u|^2 - \langle w'', u \rangle \leq -v_0|u|^2 - \langle w, u \rangle \qquad \forall u \notin N'$$

soit, $(\bar{v}_0 - v_0)|u|^2 \geq \langle w - w'', u \rangle \qquad \forall u \notin N'$

ou encore,

$$\bar{v}_0 - v_0 \geq \frac{\langle w - w'', u \rangle}{|u|^2} \qquad \forall u \notin N'$$

Or, comme nous sommes à l'extérieur d'un voisinage de l'origine, il existe un réel c tel que :

$$|u| \geq c \qquad \forall u \notin N' \qquad \text{et} \qquad c > 0$$

Par conséquent, si nous choisissons $\bar{v}_0 = \max\left(v_0'', v_0 + \frac{|w - w''|}{c}\right)$ les inégalités (1) et (2) seront vérifiées.

ANNEXE VIII

Démonstration du théorème III.2

Nous allons décomposer la démonstration en deux grandes parties, chacune correspondant à un morceau de la thèse.

I) Nous allons démontrer l'assertion suivante :
Si la topologie sur U est engendrée par une norme, si c appartenant à $\mathbb{R}^{U\times V}$ est flexible en o et si G est fermé pour l'ordre,
Alors tout point d'adhérence de la suite $\{x_k\}_{k=1}^{\infty}$, construite dans l'énoncé du théorème, satisfait les contraintes.

Dans une première étape, nous allons démontrer les propositions suivantes :

(1) $\exists\, u_k \in U$ t.q. $c(u_k, v_k) \geq \bar{c}(G(x_k), v_k) - \delta_k$ et $G(x_k) \leq u_k$

(2) $-\alpha \cdot \lambda_k \underset{(a)}{\leq} -f_o(x_k) + c(u_k, v_k) + \delta_k \underset{(b)}{\leq} -h(u_k) + c(u_k, v_k) + \delta$

où $\delta_k = \frac{1}{2}(-\eta_k - \varepsilon_k - \lambda_k)$

Remarquons d'abord que $g(v_k)$ est fini pour tout entier k.
En effet, par construction, il existe un voisinage N'_k de o tel que N'_k est inclus dans N_k et

$$\begin{cases} c(u, v_k) \leq c(u, v'_k) & \forall u \in N'_k \\ c(u, v_k) \leq c(u, v'_k) + \eta_k & \forall u \notin N'_k \end{cases} \quad (3)$$

Or, η_k est strictement négatif. Par conséquent,

$$c(u, v_k) \leq c(u, v'_k) \qquad \forall u \in U \qquad (4)$$

Comme v'_k appartient à $c\text{-}\partial_{\varepsilon_k} h(\sigma)$ c'est à dire

$$h(u) - h(\sigma) \geq c(u, v'_k) - c(\sigma, v'_k) - \varepsilon_k \qquad \forall u \in U$$

nous avons, suite à la normalisation de c en σ et à (4)

$$h(u) - h(\sigma) \geq c(u, v_k) - \varepsilon_k \qquad \forall u \in U$$

donc, $v_k \in c\text{-}\partial_{\varepsilon_k} h(\sigma)$

Ceci est valable pour tout entier k et comme $\{\varepsilon_k\}_{k=1}^{\infty}$ converge vers zéro cela implique la convergence de $\{g(v_k)\}_{k=1}^{\infty}$ vers $-\alpha$. Comme α est fini, $g(v_k)$ sera fini à partir d'un certain rang.

Nous négligerons tous les termes de la suite avant ce rang mais pour des facilités d'écriture nous garderons les mêmes notations.

Nous savons que :

$$g(v) = \sup_{x \in X} l(x, v) \qquad \forall v \in V$$

Donc, puisque $g(v_k)$ est fini pour tout entier k et comme nous travaillons avec un programme mathématique, les valeurs $\bar{c}(G(x_k), v_k)$ $(k \in \mathbb{N})$ sont également finies.

Il s'ensuit, par définition de $\bar{c}$ et de δ_k :

$\forall k \in \mathbb{N}, \exists u_k \in U, u_k \geq G(x_k)$ tel que

$$c(u_k, v_k) \geq \bar{c}(G(x_k), v_k) - \delta_k .$$

La proposition (1) est donc bien vérifiée.

Démontrons maintenant les inégalités (2)(a) et (2)(b)

(a) Nous savons par hypothèse et par la propriété (1) du lagrangien que:

$$-\alpha = \beta \leq \sup_{x \in X} [-f_0(x) + \bar{c}(G(x), v_k)] \leq l(x_k, v_k) + \lambda_k .$$

En transformant cette expression et en utilisant (1), nous obtenons l'inégalité (2)(a), soit :

$$-\alpha-\lambda_k \leq -f_0(x_k)+\bar{c}(G(x_k),v_k) \leq -f_0(x_k)+c(u_k,v_k)+\delta_k .$$

(b) Il suffit de démontrer que $-f_0(x_k) \leq -h(u_k)$

or, $h(u_k) \overset{déf}{=} \inf\limits_{\{x \mid G(x)\leq u_k\}} f_0(x)$

par (1), $G(x_k)\leq u_k$

donc, nous avons l'inégalité (2)(b).

Dans la deuxième étape, nous allons prouver (en utilisant les inégalités (1) et (2), la flexibilité de c et la construction de $\{v'_k\}_{k=1}^{\infty}$) que :

$$\forall k \in \mathbb{N}, \quad u_k \in N_k \qquad (5)$$

Supposons par l'absurde que :

$\exists k \in \mathbb{N}$ tel que $u_k \notin N_k$

donc par (3),

$u_k \notin N'_k$ et $c(u_k,v_k) \leq c(u_k,v'_k)+\eta_k$

en utilisant l'inégalité (2), nous obtenons :

$$-\alpha-\lambda_k \leq -h(u_k)+c(u_k,v'_k)+\eta_k+\delta_k$$

or, $g(v'_k) \overset{déf}{=} \sup\limits_{u\in U} \{c(u,v'_k)-h(u)\}$ (6)

par conséquent,

$$-\alpha-\lambda_k \leq g(v'_k)+\delta_k+\eta_k$$

en remplaçant δ_k par sa valeur, nous avons :

$$-\alpha-\lambda_k \leq g(v'_k)+\tfrac{1}{2}\eta_k-\tfrac{1}{2}\varepsilon_k+\tfrac{1}{2}\lambda_k$$

Comme $\eta_k < -\varepsilon_k-\lambda_k$, il s'ensuit

$$-\alpha < g(v'_k)-\varepsilon_k \qquad (7)$$

Par hypothèse, nous savons que:

$$h(u)-h(\sigma) \geq c(u,v'_k)-c(\sigma,v'_k)-\varepsilon_k \qquad \forall u\in U$$

ou, $$-\alpha \geq c(u,v'_k)-h(u)-\varepsilon_k \qquad \forall u \in U$$

en utilisant (6) nous obtenons:

$$-\alpha \geq g(v'_k)-\varepsilon_k \qquad (8)$$

(7) et (8) sont en contradiction donc l'hypothèse de départ est fausse et (5) est bien vérifié.

En utilisant les faits suivants:

. U est un espace vectoriel normé (9)

. G fermé pour l'ordre (10)

. $G(x_k) \leq u_k$ et $u_k \in N_k$ $\forall k \in \mathbb{N}$ (11)

nous allons démontrer dans la troisième étape, que tout point d'adhérence de $\{x_k\}_{k=1}^{\infty}$ satisfait les contraintes.

En effet, comme $\{N_k\}_{k=1}^{\infty}$ est une suite de boules ouvertes centrées en σ dont les diamètres décroissent vers zéro, nous avons par (9) et (11) que la suite $\{u_k\}_{k=1}^{\infty}$ converge vers σ.

Soit x^* un point d'adhérence de $\{x_k\}_{k=1}^{\infty}$, par (10) et (11) nous obtenons:

$$G(x^*) \leq \sigma$$

Par conséquent x^* satisfait les contraintes.

II) La deuxième partie de la démonstration consistera à prouver l'assertion suivante:

Si, en plus des hypothèses précédentes, f_0 est s.c.i, $\{\eta_k\}_{k=1}^{\infty}$ et $\{\lambda_k\}_{k=1}^{\infty}$ convergent vers zéro et si la famille $\{c(\cdot,v_k)\}_{k=1}^{\infty}$ est s.c.s en σ uniformément par rapport à k, alors tout point d'adhérence de $\{x_k\}_{k=1}^{\infty}$ est une solution du problème initial.

Dans une première étape, nous démontrons que :

$$\limsup_{k \to \infty} c(u_k, v_k) \leq 0 \qquad (12)$$

Soit ε, un réel strictement positif arbitraire.
Vu l'hypothèse sur la famille $\{c(\cdot, v_k)\}_{k=1}^{\infty}$, nous savons qu'il existe un voisinage V de σ tel que :

$$\forall k \in \mathbb{N}, \ \forall u \in V \quad c(u, v_k) \leq c(\sigma, v_k) + \varepsilon.$$

En outre, comme $\{u_k\}_{k=1}^{\infty}$ converge vers σ il existe un rang $K \in \mathbb{N}$ tel que : $\forall k \geq K \quad u_k \in V$

donc, $\forall k \geq K \quad c(u_k, v_k) \leq \varepsilon$

ε étant arbitraire, nous en déduisons l'inégalité (12).

Dans la deuxième étape, nous allons enfin conclure que tout point d'adhérence de $\{x_k\}_{k=1}^{\infty}$ est une solution du problème initial.

Soit x^* un point d'adhérence de $\{x_k\}_{k=1}^{\infty}$. Notons de la même façon la sous suite de $\{x_k\}$ convergeant vers x^*.

Par (2)(a) :

$$-\alpha - \lambda_k \leq -f_0(x_k) + c(u_k, v_k) + \delta_k \qquad \forall k \in \mathbb{N}.$$

en prenant la limite supérieure des deux membres pour $k \to \infty$ nous obtenons :

$$-\alpha \leq \limsup_{k \to \infty} \left(-f_0(x_k) + c(u_k, v_k) + \delta_k\right)$$

donc, par définition de δ_k et par (12) :

$$-\alpha \leq -\liminf_{k \to \infty} f_0(x_k) = -f_0(x^*) \quad \text{car } f_0 \text{ est s.c.i}$$

Par conséquent, $\inf_{\{x \in X \mid G(x) \leq \sigma\}} f_0(x) = \alpha \geq f_0(x^*)$

or, par I), x^* satisfait les contraintes donc $f_0(x^*) = \alpha$ et x^* est bien une solution optimale de (P).

Démonstration du théorème de Rockafellar

Dans une première étape, nous allons démontrer que :
Sous les hypothèses du théorème, $\{x_k\}_{k=1}^{\infty}$ satisfait asymptotiquement les contraintes et $\liminf_{k\to\infty} \frac{w_{ik}}{v_{0k}} \geq 0 \qquad i = 1..m$

Par hypothèse, $l(x_k, v_k) \geq \sup_{x\in X} l(x, v_k) - \alpha_k$

Or, $\sup_{x\in X} l(x, v_k) = g(v_k)$ et $\inf_{v\in V} g(v) = \beta = -\alpha$

donc, $l(x_k, v_k) \geq g(v_k) - \alpha_k \geq \beta - \alpha_k \qquad (1)$

Nous savons également (cfr. annexe VI) que :

$$l(x_k, v_k) = -f_0(x_k) - \langle w_k, u_k\rangle - v_{0k}|u_k|^2 \qquad (2)$$

où $u_k = (u_{1k}, .. u_{mk}) \qquad u_{ik} \stackrel{\text{déf}}{=} \max\left(y_i(x), -\frac{w_{ik}}{2v_{0k}}\right) \quad i = 1,..m$

Comme $f_0(x_k) \geq h(u_k) = \inf\{f_0(x) \mid x \in X, G(x) \leq u_k\}$

nous obtenons :

$$l(x_k, v_k) \leq -h(u_k) - \langle w_k, u_k\rangle - (v_{0k} - \delta)|u_k|^2 - \delta|u_k|^2.$$

et donc, en utilisant la définition de g et de v'_k, nous avons :

$$l(x_k, v_k) \leq g(v'_k) - \delta|u_k|^2 \qquad (3)$$

Réunissons les résultats (1) et (3) il s'ensuit :

$$\beta - \alpha_k \leq g(v'_k) - \delta|u_k|^2$$

soit, $\delta|u_k|^2 \leq g(v'_k) - \beta + \alpha_k$.

Prenons la limite des deux membres pour $k \to \infty$, cela implique :

$$\lim_{k\to\infty} \delta|u_k|^2 = 0 \qquad \text{car } \delta|u_k|^2 \geq 0 \quad \forall k \in N \qquad (4)$$

Puisque $\delta > 0$ et par définition de la norme, nous avons donc :

$$\lim_{k\to\infty} u_{ik} = 0 \qquad i = 1,..m$$

Donc, en utilisant la définition de u_k, nous obtenons :

1) $\limsup_{k \to \infty} g_i(x_k) \leq 0 \qquad i = 1, \dots m$

Par conséquent $\{x_k\}_{k=1}^{\infty}$ satisfait asymptotiquement les contraintes.

2) $\limsup_{k \to \infty} \left(-\frac{w_{ik}}{2v_{0k}}\right) \leq 0 \qquad i = 1, \dots m$

Remarque : Si nous ajoutons l'hypothèse "G fermé pour l'ordre" alors un point d'adhérence x^* quelconque de $\{x_k\}_{k=1}^{\infty}$ satisfait les contraintes.

En effet, $G(x_k) \leq u_k \quad \forall k \in \mathbb{N}$.

et $\{u_k\}_{k=1}^{\infty}$ converge vers $\sigma = (0, \dots 0) \in \mathbb{R}^m$.

Par conséquent, puisque G fermé pour l'ordre, $G(x^*) \leq \sigma$.

Dans une deuxième étape, nous allons prouver que si, en outre, $\{w_k\}_{k=1}^{\infty}$ est une suite bornée alors $\{x_k\}_{k=1}^{\infty}$ minimise asymptotiquement (P).

Par (1) et (3) nous avons :

$$\beta - \alpha_k \leq l(x_k, v_k) \leq g(v'_k) - \delta |u_k|^2.$$

En utilisant (4) et les hypothèses de convergence sur les suites $\{g(v'_k)\}_{k=1}^{\infty}$ et $\{\alpha_k\}_{k=1}^{\infty}$ nous concluons, par passage à la limite :

$$\lim_{k \to \infty} l(x_k, v_k) = \beta$$

Comme $\{w_k\}_{k=1}^{\infty}$ est une suite bornée, (2) et (4) impliquent :

$$\lim_{k \to \infty} -f_0(x_k) = \lim_{k \to \infty} l(x_k, v_k) = \beta = -\alpha.$$

Remarque : Si nous ajoutons l'hypothèse "f_0 s.c.i" alors un point d'ahérence x^* quelconque de $\{x_k\}_{k=1}^{\infty}$ sera une solution optimale de (P).

En effet, sélectionnons la sous suite de $\{x_k\}_{k=1}^{\infty}$ convergeant vers x^*, et pour plus de facilités notons-la toujours par $\{x_k\}_{k=1}^{\infty}$, nous avons :

$$\alpha = \lim_{k \to \infty} f_0(x_k) = \liminf_{k \to \infty} f_0(x_k) = f_0(x^*).$$

ANNEXE IX

Démonstration de la proposition III.2

Nous allons démontrer que :

$\forall \varepsilon > 0$, $\exists V_\varepsilon$ un voisinage de σ tel que :

$\forall i \in I$, $\forall u \in V_\varepsilon \qquad a_i(u) \leq a_i(\sigma) + \varepsilon$.

Par hypothèse, a_i est concave donc la dérivée directionnelle de a_i en σ existe dans $\overline{\mathbb{R}}$ dans toutes les directions et en outre,

$$a_i'(\sigma\,;u) = \sup_{\lambda > 0} \frac{a_i(\sigma + \lambda u) - a_i(\sigma)}{\lambda} \qquad \forall u \in U, \forall i \in I \qquad (1)$$

En effet, $\dfrac{a_i(\sigma+\lambda u) - a_i(\sigma)}{\lambda}$ est une fonction décroissante en $\lambda > 0$

$\forall i \in I \qquad (2)$

Soient u appartenant à U et deux réels λ_1, λ_2 tels que : $0 < \lambda_1 < \lambda_2$

$$\sigma + \lambda_1 u = \frac{\lambda_1}{\lambda_2}(\sigma + \lambda_2 u) + \left(1 - \frac{\lambda_1}{\lambda_2}\right)\sigma$$

Soit un élément i arbitraire dans I.

Comme a_i est concave, nous avons :

$$a_i(\sigma + \lambda_1 u) \geq \frac{\lambda_1}{\lambda_2} a_i(\sigma + \lambda_2 u) + \left(1 - \frac{\lambda_1}{\lambda_2}\right) a_i(\sigma)$$

donc,

$$\frac{a_i(\sigma + \lambda_1 u) - a_i(\sigma)}{\lambda_1} \geq \frac{a_i(\sigma + \lambda_2 u) - a_i(\sigma)}{\lambda_2}$$

L'assertion (2) et par conséquent la (1) sont donc bien vérifiées.

Nous savons par hypothèse que : $a_i'(\sigma\,;u) \leq K \quad \forall i \in I$, $\forall u \in U$ t.q $\|u\| = 1$.

Donc, par (1) et en utilisant le fait que σ est le neutre pour l'addition dans U, il s'ensuit :

$\forall \lambda > 0$, $\forall u \in U$ tel que $\|u\| = 1$ $\quad a_i(\lambda u) \leq a_i(\sigma) + \lambda K \quad (3)$

Soit ε un réel arbitraire, $\varepsilon > 0$, posons $V_\varepsilon^o = B(\sigma, \frac{\varepsilon}{K})$ (boule ouverte de centre σ et de rayon $\frac{\varepsilon}{K}$)

Tout élément u de V_ε^o peut s'écrire $\lambda u'$ où $\|u'\| = 1$ et $\lambda < \frac{\varepsilon}{K}$

Par (3), nous avons donc :

$\forall i \in I$, $\forall u \in V_\varepsilon^o$ $\quad a_i(u) \leq a_i(\sigma) + \lambda K \leq a_i(\sigma) + \varepsilon$

Par conséquent, nous avons bien la thèse.

Nous allons démontrer que la famille $\{c(\cdot, v)\}_{v \in V}$ où c désigne la fonctionnelle couplante de l'exemple 1.a est semi-continue supérieurement en σ uniformément par rapport à v.

Si nous considérons la fonctionnelle couplante de l'exemple 2.a nous avons que toute famille du type $\{c(\cdot, v_i)\}_{i \in I}$ (où $v_i = (v_{0i}, v_i)$ et $\|v_i\| \leq K$ $\forall i \in I$, $K \in \mathbb{R}$) est semi-continue supérieurement en σ uniformément par rapport à i

<u>Exemple 1.a</u>

$U = \mathbb{R}^m$, $V = \mathbb{R}_+^m$

$c(u, v) = -\sum_{i=1}^{m} v_i |u_i| \qquad u = (u_1, \dots u_m) \in \mathbb{R}^m$ et $v = (v_1, \dots v_m) \in \mathbb{R}_+^m$

l'assertion énoncée est immédiate.

En effet, $\forall \varepsilon > 0$ $\exists V_\varepsilon^o$ un voisinage de $\sigma = (0, 0 \dots 0) \in \mathbb{R}^m$ tel que :

$c(u, v) \leq c(\sigma, v) + \varepsilon \quad \forall v \in V, \forall u \in V_\varepsilon^o$ (car $\forall u \in U, \forall v \in V$ $c(u, v) \leq 0$)

Exemple 2.a

. $U = H$ un espace de Hilbert, $V = \mathbb{R}_+ \times H$, H est supposé de dimension finie.

. $c(u, v) = -v_0 |u|^2 - \langle w, u \rangle$, $v = (v_0, w) \in \mathbb{R}_+ \times H$, $u \in H$

Nous allons démontrer que les hypothèses de la proposition III.2 sont satisfaites pour une famille arbitraire $\{c(\cdot, v_i)\}_{i \in I}$ satisfaisant : $\forall i \in I \quad v_i = (v_{0_i}, w_i)$ et $\| w_i \| \leq K$, $K \in \mathbb{R}$.

1) Soit un élément v de V arbitraire

Considérons un réel λ quelconque appartenant à $]0,1[$ et deux éléments u_1, u_2 choisis arbitrairement dans U.
Nous devons prouver que $c(\cdot, v)$ est concave c'est à dire :

$$c(\lambda u_1 + (1-\lambda) u_2, v) \leq \lambda c(u_1, v) + (1-\lambda) c(u_2, v)$$

En transformant cette expression en tenant compte de la définition de c, cela revient à vérifier l'inégalité suivante qui est immédiate.

$$\underbrace{\lambda(\lambda-1)}_{<0} \underbrace{[|u_1|^2 + |u_2|^2 - 2|u_1|.|u_2|]}_{\geq 0} \leq 0$$

Donc $c(\cdot, v)$ est concave pour tout v dans V.

2) Calculons la dérivée directionnelle de $c(\cdot, v)$ en σ dans une direction u arbitraire (nous noterons $c_v(\cdot)$ à la place de $c(\cdot, v)$ pour plus de facilités.)

$$c'_v(\sigma; u) \overset{\text{déf}}{=} \lim_{\lambda \to 0^+} \frac{-v_0 |\sigma + \lambda u|^2 - \langle w, \sigma + \lambda u \rangle + v_0 |\sigma|^2 + \langle w, \sigma \rangle}{\lambda}$$

Comme σ est le neutre pour l'addition dans U, nous avons :

$$c'_v(\sigma; u) = \lim_{\lambda \to 0^+} \frac{-\lambda^2 v_0 |u|^2 - \lambda \langle w, u\rangle}{\lambda}$$

$$c'_v(\sigma; u) = -\langle w, u\rangle \qquad \forall u \in U$$

Donc par l'inégalité de Cauchy-Schwartz, nous obtenons :

$c'_v(\sigma; u) \leq \|w\| . \|u\| \qquad \forall u \in U$

Par conséquent, si nous nous limitons aux éléments de norme 1 dans U,

$c'_v(\sigma; u) \leq \|w\| \qquad \forall u \in U$ tel que $\|u\| = 1$.

D'où, si nous considérons une famille $\{c(\cdot, v_i)\}_{i \in I}$ où $v_i = (v_{0_i}, w_i)$ et $\|w_i\| \leq K \quad \forall i \in I$ (1) (K étant une constante réelle) toutes les hypothèses de la proposition III.2 seront bien vérifiées.

Remarquons que la condition (1) n'est pas tellement restrictive car on peut encore faire varier le paramètre v_0 comme on veut.

ANNEXE X

Démonstration du théorème III.3

Dans une première étape nous allons vérifier l'implication suivante :

$$\bar{v} \in c\text{-}\partial h(\sigma) \implies \bar{\bar{v}} \in c\text{-}\partial h(\sigma) \qquad (1)$$

Par hypothèse, il existe N' un voisinage de σ inclus dans N, tel que :

$$\begin{cases} c(u,\bar{\bar{v}}) \le c(u,\bar{v}) & \forall u \in N' \quad (2) \\ c(u,\bar{\bar{v}}) \le c(u,\bar{v}) + \eta & \forall u \notin N' \quad (3) \end{cases}$$

Comme η est strictement négatif, nous obtenons :

$$c(u,\bar{\bar{v}}) \le c(u,\bar{v}) \qquad \forall u \in U \qquad (4)$$

Or $\bar{v} \in c\text{-}\partial h(\sigma)$ est équivalent à l'inégalité :

$$h(u) - h(\sigma) \ge c(u,\bar{v}) - c(\sigma,\bar{v}) \qquad \forall u \in U$$

donc par (4) et par la normalisation de c en σ nous déduisons :

$$h(u) - h(\sigma) \ge c(u,\bar{\bar{v}}) - c(\sigma,\bar{\bar{v}}) \qquad \forall u \in U$$

Par conséquent,

$$\bar{\bar{v}} \in c\text{-}\partial h(\sigma)$$

Dans la seconde étape nous vérifions que $\tilde{x}$ est une solution optimale de (P).

Par hypothèse, il existe un élément u_1 de U tel que :

$$G(\tilde{x}) \le u_1 \quad \text{et} \quad \sup\{c(u,\bar{\bar{v}}) \mid u \in U, G(\tilde{x}) \le u\} = c(u_1,\bar{\bar{v}})$$

Or, $\sup_{x\in X} l(x,\bar{v}) = l(\tilde{x},\bar{v})$ et $\bar{\bar{v}} \in c\text{-}\partial h(\sigma)$ ce qui implique :

$$g(\bar{\bar{v}}) = \sup_{x\in X} l(x,\bar{v}) = -\alpha = \beta$$

donc, par définition de h et de u_1 nous obtenons:

$$-\alpha = l(\tilde{x}, \bar{\bar{v}}) \overset{\text{déf}}{=} -f_0(\tilde{x}) + \bar{c}(G(\tilde{x}), \bar{\bar{v}}) \leq -h(u_1) + c(u_1, \bar{\bar{v}})$$

Par conséquent, l'inégalité suivante est vérifiée.

$$-\alpha \leq (-h(u_1) + c(u_1, \bar{v})) + c(u_1, \bar{\bar{v}}) - c(u_1, \bar{v}) \qquad (5)$$

Nous savons également que $\bar{v}$ appartient à $c\text{-}\partial h(\sigma)$ c'est à dire : $h(u) - h(\sigma) \geq c(u, \bar{v}) - c(\sigma, \bar{v}) \qquad \forall u \in U$

Ce qui entraîne, en utilisant la normalisation de c en σ

$$-\alpha \geq -h(u_1) + c(u_1, \bar{v}) \qquad (6)$$

En réunissant les résultats (5) et (6) nous obtenons:

$$c(u_1, \bar{v}) \leq c(u_1, \bar{\bar{v}}) \qquad (7)$$

par conséquent, suite à (2) et (3) et comme η est strictement négatif, u_1 doit appartenir à N', donc à N.

Or, par hypothèse $c(u, \bar{\bar{v}}) < c(u, \bar{v}) \quad \forall u \in N \setminus \{\sigma\}$

donc, $u_1 = \sigma$

Il s'ensuit : $G(\tilde{x}) \leq \sigma$ c'est à dire $\tilde{x}$ satisfait les contraintes et $\bar{c}(G(\tilde{x}), \bar{\bar{v}}) = c(u_1, \bar{\bar{v}}) = 0$

Ce qui implique $-\alpha = l(\tilde{x}, \bar{\bar{v}}) = -f_0(\tilde{x})$

Donc, $\tilde{x}$ est bien une solution optimale de (P).

RÉFÉRENCES

[1] E.J. BALDER (1976) "An extension of duality - satibility relations to nonconvex optimization problems". SIAM J. Control and Optimization. Vol 15, No 2, February 1976.

[2] P.J. LAURENT (1972) "Approximation et optimisation" Hermann, Paris.

[3] I. EKELAND, R. TEMAM (1973) "Analyse convexe et problèmes variationnels". Dunod, Gauthier - Villars collection: études mathématiques.

[4] P.O. LINDBERG (1976) "A generalization of Fenchel conjugation giving generalized lagrangians and symmetric nonconvex duality" Optimization and systems theory. TRITA-MAT-1976-12 (Aug).

[5] LUENBERGER (1973) "Introduction to linear and non linear programming". Addison - Wesley.

[6] D.P. BERTSEKAS (1976) "Multiplier methods: a survey". Automatica, Vol 12, pp 133-145. Pergamon Press.

[7] R.T. ROCKAFELLAR (1974) "Augmented Lagrange Multiplier Functions and Duality in nonconvex programming". SIAM J. Control, Vol 12, No 2, May 1974.

[8] R.T. ROCKAFELLAR (1970) "Convex Analysis."
Princeton.

Table des matières